

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 September 2002 (06.09.2002)

PCT

(10) International Publication Number
WO 02/068600 A2

- (51) International Patent Classification⁷: C12N (74) Agents: STRAHER, Michael, P. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, 46th floor, One Liberty Place, Philadelphia, PA 19103 (US).
- (21) International Application Number: PCT/US02/05625
- (22) International Filing Date: 26 February 2002 (26.02.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/271,913 26 February 2001 (26.02.2001) US
- (71) Applicant (for all designated States except US): ARENA PHARMACEUTICALS, INC. [US/US]; 6166 Nancy Drive, San Diego, CA 92121 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): LIAW, Chen, W. [US/US]; 7668 Salix Place, San Diego, CA 92129 (US). CHALMERS, Derek, T. [GB/US]; 347 Longden Lane, Solana Beach, CA 92075 (US). BEHAN, Dominic, P. [GB/US]; 11472 Roxboro Court, San Diego, CA 92131 (US). MACIEJEWSKI-LENIOR, Dominique [US/US]; 3615 Luna Avenue, San Diego, CA 92117 (US). LEONARD, James, N. [US/US]; 11326 Via Playa de Cortes, San Diego, CA 92124 (US). LIN, I-Lin [—/US]; 8291-7 Gold Coast Drive, San Diego, CA 92126 (US). ORTUNO, Daniel [US/US]; 1233 Adobe Terrace, Vista, CA 92083 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENDOGENOUS AND NON-ENDOGENOUS VERSIONS OF HUMAN G PROTEIN-COUPLED RECEPTORS

(57) Abstract: The invention disclosed in this patent document relates to transmembrane receptors, more particularly to a human G protein-coupled receptor and to mutated (non-endogenous) versions of the human GPCRs for evidence activity.



WO 02/068600 A2

**ENDOGENOUS AND NON-ENDOGENOUS VERSIONS OF
HUMAN G PROTEIN-COUPLED RECEPTORS**

5

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Serial Number 09/170,496, filed on October 13, 1998 and its corresponding PCT application number PCT/US99/23938, published as WO 00/22129 on April 20, 2000. This application also is a continuation in part of U.S. Ser. No. 09/060,188, filed April 14, 1998, which is a continuation in part of U.S. Ser. No. 08/839,449, filed April 14, 1997 (abandoned). The priority benefit of each of the foregoing is claimed herein, and the disclosures of each of the foregoing is incorporated by reference herein in its entirety. This application also claims the benefit of U.S. Provisional Number 60/271,913, filed February 26, 2001, also incorporated herein by reference in its entirety. This document is related to the following applications: U.S. Provisional Number 60/250,881, filed December 1, 2000; U.S. Provisional Number 60/253,428, filed November 27, 2000; U.S. Provisional Number 60/234,317, filed September 20, 2000; U.S. Provisional Number 60/245,853, filed November 3, 2000; U.S. Provisional Number 60/234,045, filed September 20, 2000; U.S. Provisional Number 60/200,568, filed April 28, 2000; U.S. Provisional Number 60/198,518, filed April 19, 2000; U.S. Provisional Number 60/189,353, filed March 14, 2000; U.S. Provisional Number 60/166,084, filed November 17, 1999; and U.S. Provisional Number 60/106,451, filed October 30, 1998, the disclosures of each of which are incorporated herein by reference in their entirety.

FIELD OF THE INVENTION

The present invention relates to transmembrane receptors, in some embodiments to G protein-coupled receptors and, in some preferred embodiments, to endogenous GPCRs that are altered to establish or enhance constitutive activity of the receptor. In some 5 embodiments, the constitutively activated GPCRs will be used for the direct identification of candidate compounds as receptor agonists or inverse agonists having applicability as therapeutic agents.

10 BACKGROUND OF THE INVENTION

Although a number of receptor classes exist in humans, by far the most abundant and therapeutically relevant is represented by the G protein-coupled receptor (GPCR) class. It is estimated that there are some 30,000-40,000 genes within the human genome, and of 15 these, approximately 2% are estimated to code for GPCRs. Receptors, including GPCRs, for which the endogenous ligand has been identified, are referred to as "known" receptors, while receptors for which the endogenous ligand has not been identified are referred to as "orphan" receptors.

GPCRs represent an important area for the development of pharmaceutical 20 products: from approximately 20 of the 100 known GPCRs, approximately 60% of all prescription pharmaceuticals have been developed. For example, in 1999, of the top 100 brand name prescription drugs, the following drugs interact with GPCRs (diseases and/or disorders treated are indicated in parentheses):

Claritin® (allergies)	Prozac® (depression)	Vasotec® (hypertension)
25 Paxil® (depression)	Zoloft® (depression)	Zyprexa® (psychotic disorder)
Cozaar® (hypertension)	Imitrex® (migraine)	Zantac® (reflux)
Propulsid® (reflux disease)	Risperdal® (schizophrenia)	Serevent® (asthma)
Pepcid® (reflux)	Gaster® (ulcers)	Atrovent® (bronchospasm)

Effexor® (depression)	Depakote® (epilepsy)	Cardura® (prostatic hypertrophy)
Allegra® (allergies)	Lupron® (prostate cancer)	Zoladex® (prostate cancer)
Diprivan® (anesthesia)	BuSpar® (anxiety)	Ventolin® (bronchospasm)
Hytrin® (hypertension)	Wellbutrin® (depression)	Zyrtec® (rhinitis)
5 Plavix® (MI/stroke)	Toprol-XL® (hypertension)	Tenomin® (angina)
Xalatan® (glaucoma)	Singulair® (asthma)	Diovan® (hypertension)
Harnal® (prostatic hyperplasia)		
(Med Ad News 1999 Data).		

GPCRs share a common structural motif, having seven sequences of between 22 to 24 hydrophobic amino acids that form seven alpha helices, each of which spans the membrane (each span is identified by number, *i.e.*, transmembrane-1 (TM-1), transmembrane-2 (TM-2), *etc.*). The transmembrane helices are joined by strands of amino acids between transmembrane-2 and transmembrane-3, transmembrane-4 and transmembrane-5, and transmembrane-6 and transmembrane-7 on the exterior, or “extracellular” side, of the cell membrane (these are referred to as “extracellular” regions 1, 2 and 3 (EC-1, EC-2 and EC-3), respectively). The transmembrane helices are also joined by strands of amino acids between transmembrane-1 and transmembrane-2, transmembrane-3 and transmembrane-4, and transmembrane-5 and transmembrane-6 on the interior, or “intracellular” side, of the cell membrane (these are referred to as “intracellular” regions 1, 2 and 3 (IC-1, IC-2 and IC-3), respectively). The “carboxy” (“C”) terminus of the receptor lies in the intracellular space within the cell, and the “amino” (“N”) terminus of the receptor lies in the extracellular space outside of the cell.

Generally, when an endogenous ligand binds with the receptor (often referred to as “activation” of the receptor), there is a change in the conformation of the intracellular region that allows for coupling between the intracellular region and an intracellular “G-protein.” It

has been reported that GPCRs are "promiscuous" with respect to G proteins, *i.e.*, that a GPCR can interact with more than one G protein. *See*, Kenakin, T., 43 *Life Sciences* 1095 (1988). Although other G proteins exist, currently, G_q, G_s, G_i, G_z and G_o are G proteins that have been identified. Ligand-activated GPCR coupling with the G-protein initiates a signaling cascade process (referred to as "signal transduction"). Under normal conditions, signal transduction ultimately results in cellular activation or cellular inhibition. Although not wishing to be bound to theory, it is thought that the IC-3 loop as well as the carboxy terminus of the receptor interact with the G protein.

Under physiological conditions, GPCRs exist in the cell membrane in equilibrium between two different conformations: an "inactive" state and an "active" state. A receptor in an inactive state is unable to link to the intracellular signaling transduction pathway to initiate signal transduction leading to a biological response. Changing the receptor conformation to the active state allows linkage to the transduction pathway (via the G-protein) and produces a biological response.

A receptor may be stabilized in an active state by a ligand or a compound such as a drug. Recent discoveries, including but not exclusively limited to modifications to the amino acid sequence of the receptor, provide means other than ligands or drugs to promote and stabilize the receptor in the active state conformation. These means effectively stabilize the receptor in an active state by simulating the effect of a ligand binding to the receptor. Stabilization by such ligand-independent means is termed "constitutive receptor activation."

SUMMARY OF THE INVENTION

Disclosed herein are endogenous and non-endogenous versions of human GPCRs and uses thereof.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:2, non-endogenous, constitutively activated versions of the same encoded by an amino acid of SEQ.ID.NO.:63, and host cells comprising the same.

5 Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:62 and host cells comprising the same.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:4, non-endogenous, constitutively activated versions of the same encoded by an amino acid of
10 SEQ.ID.NO.:65, and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:64 and host cells comprising the same.

Some embodiments of the present invention relate to G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:6, non-endogenous, constitutively
15 activated versions of the same, and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:5 and host cells comprising the same.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:8, non-endogenous,
20 constitutively activated versions of the same encoded by an amino acid of SEQ.ID.NO.:67, SEQ.ID.NO.:69, SEQ.ID.NO.:71, and SEQ.ID.NO.:73, and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:66, SEQ.ID.NO.:68, SEQ.ID.NO.:70, and SEQ.ID.NO.:72, and host cells comprising the same.

Some embodiments of the present invention relate to a G protein-coupled
5 receptor encoded by an amino acid sequence of SEQ.ID.NO.:10, non-endogenous, constitutively activated versions of the same encoded by an amino acid of SEQ.ID.NO.:75 and SEQ.ID.NO.:77, and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:74 and SEQ.ID.NO.:76, and host cells comprising
10 the same.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:12, non-endogenous, constitutively activated versions of the same encoded by an amino acid of SEQ.ID.NO.:79 and SEQ.ID.NO.:81, and host cells comprising the same.

15 Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:78 and SEQ.ID.NO.:80, and host cells comprising the same.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:14, constitutively activated
20 versions of the same encoded by an amino acid of SEQ.ID.NO.:83, and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:82 and host cells comprising the same.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:16, constitutively activated versions of the same encoded by an amino acid of SEQ.ID.NO.:85, and host cells comprising the same.

5 Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:84 and host cells comprising the same.

Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:18, constitutively activated versions of the same encoded by an amino acid of SEQ.ID.NO.:87, and host cells
10 comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:86 and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:84 and host cells comprising the same.

15 Some embodiments of the present invention relate to a G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:98, non-endogenous, constitutively activated versions of the same and host cells comprising the same.

Some embodiments of the present invention relate to a plasmid comprising a vector and the cDNA of SEQ.ID.NO.:97 and host cells comprising the same.

20

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a graphic representation of the results of a second messenger cell-based cyclic AMP assay providing comparative results for constitutive signaling of endogenous, constitutively active FPRL-2 ("FPRL-2 wt"), non-endogenous, constitutively activated

version of FPRL-2 ("FPRL-2 (L240K)") fused with a Gs/Gi Fusion Protein Construct and a control ("Gs/Gi").

Figure 2 provides graphic results of comparative analysis of endogenous STRL33 against non-endogenous, constitutively activated STRL33 ("STRL33(L230K)") utilizing an 8XCRE-Luc Reporter assay in 293T cells as compared with the control ("CMV").

Figure 3 provides graphic results of comparative analysis of a co-transfection of non-endogenous TSHR(A623I) ("signal enhancer") with an endogenous target receptor, in this case GPR45 ("GPR45 wt"), versus a control ("CMV"), utilizing a cell-based adenylyl cyclase assay in 293 cells. This assay involved the addition of TSH, the endogenous ligand for TSHR.

Figure 4 provides graphic results of comparative analysis of a co-transfection of non-endogenous TSHR(A623I) ("signal enhancer") and an endogenous target receptor, in this case mGluR7 ("mGluR7 wt"), versus non-endogenous, constitutively activated versions of the target receptor mGluR7 ("W590S," "R659H" "T771C" and "T790K") co-transfected with non-endogenous TSHR(A623I), utilizing a cell-based adenylyl cyclase assay in 293 cells. This assay involved the addition of TSH, the endogenous ligand for TSHR.

Figure 5 provides graphic results of comparative analysis of a co-transfection of non-endogenous TSHR(A623I) ("signal enhancer") and an endogenous target receptor, in this case mGluR7 ("mGluR7 wt"), versus non-endogenous, constitutively activated versions of the target receptor mGluR7 ("W590S," "R659H" "T771C" and "T790K") co-transfected with non-endogenous TSHR(A623I), utilizing a cell-based adenylyl cyclase assay in RGT cells. This assay involved the addition of TSH, the endogenous ligand for TSHR.

Figure 6 provides an illustration of second messenger IP_3 production of non-endogenous mGluR7, "T771C", co-transfected with non-endogenous versions of Gq

protein, "Gq(del)" and "Gq(del)/Gi" compared with "Gq(del)" and "Gq(del)/Gi" in the presence and absence of glutamate.

Figure 7 is a comparative analysis of endogenous, non-constitutively active GPR37 ("wt") and non-endogenous, constitutively activated versions of GPR37 ("C543Y" and "L352R") in an SRE Reporter assay, where the control is expression vector ("CMV").

Figure 8 is comparative analysis of a co-transfection of Gs/Gi Fusion Construct and an endogenous target receptor, in this case GPR37 ("GPR37 wt"), versus non-endogenous, constitutively activated versions of the target receptor GPR37 ("C543Y" and "L352R") co-transfected with Gs/Gi Fusion Construct utilizing a whole cell second messenger cAMP assay.

Figure 9 is a representation of a Northern Analysis of GPR37 expressed in forskolin treated rat Schwann cells. Cell differentiation was maintained at 20uM of forskolin.

Figure 10 is a representation of a Northern Analysis of GPR37 expressed in crushed rat sciatic nerve. GPR37 was highly up-regulated seven (7) days post crush.

Figure 11 is a comparative analysis of endogenous, non-constitutively active HF1948 ("wt") and non-endogenous, constitutively activated version of HF1948 ("I281F") in an IP3 assay, where the control is expression vector ("pCMV").

Figure 12 is comparative analysis of a co-transfection of non-endogenous TSHR-A623I ("signal enhancer") and an endogenous target receptor, in this case HF1948 ("HF1948 wt"), versus non-endogenous, constitutively activated versions of the target receptor HF1948 ("I281F" and "E135N") co-transfected with non-endogenous TSHR-A623I, utilizing a whole cell adenylyl cyclase assay. This assay involved the addition of TSH, the endogenous ligand for TSHR.

Figure 13 a reproduction of a photograph of the results for the Northern Blot of GPR66 using multiple pancreatic cell lines.

Figure 14 provides graphic results of comparative analysis of endogenous GPR35 against non-endogenous, constitutively activated GPR35 ("GPR35(A216K)") utilizing an
5 E2F-Luc Reporter assay in 293A cells.

Figure 15 is a reproduction of a photograph of the results for the Northern Blot of GPR35 using multiple tissue (human) cDNA.

Figures 16 provides graphic results of comparative analysis of a co-transfection of non-endogenous TSHR-A623I ("TSHR-A623I") (with and without TSH) and endogenous
10 ETBR-LP2 ("WT"), versus non-endogenous, constitutively activated ETBR-LP2 ("N358K") co-transfected with mutated TSHR-A623I (with and without TSH) utilizing an adenylyl cyclase assay.

Figure 17 provides a graphic result comparative analysis of endogenous ETBR-LP2 ("WT") and non-endogenous, constitutively activated ETBR-LP2 ("N358K") utilizing an
15 AP1 reporter assay system.

Figure 18 is a representation of a Northern Analysis of ETBR-LP2 expressed in forskolin treated rat Schwann cells. Cell differentiation was maintained at 20uM of forskolin.

Figure 19 is a representation of a Northern Analysis of ETBR-LP2 expressed in
20 crushed rat sciatic nerve. ETBR-LP2 was highly up-regulated seven (7) days post crush.

Figures 20A and 20B provides an alignment report between the putative amino acid sequence of the human ETBR-LP2 ("hETBRLP2p") and the reported amino acid sequence of human GPR37 ("hGPR37p").

DETAILED DESCRIPTION

The scientific literature that has evolved around receptors has adopted a number of terms to refer to ligands having various effects on receptors. For clarity and consistency, the following definitions will be used throughout this patent document. To the extent that these definitions conflict with other definitions for these terms, the following definitions shall control:

AGONISTS shall mean materials (e.g., ligands, candidate compounds) that activate the intracellular response when they bind to the receptor, or enhance GTP binding to membranes. In some embodiments, **AGONISTS** are those materials not previously known to activate the intracellular response when they bind to the receptor or to enhance GTP binding to membranes.

AMINO ACID ABBREVIATIONS used herein are set out in Table A:

TABLE A

ALANINE	ALA	A
ARGININE	ARG	R
ASPARAGINE	ASN	N
ASPARTIC ACID	ASP	D
CYSTEINE	CYS	C
GLUTAMIC ACID	GLU	E
GLUTAMINE	GLN	Q
GLYCINE	GLY	G
HISTIDINE	HIS	H
ISOLEUCINE	ILE	I
LEUCINE	LEU	L
LYSINE	LYS	K

METHIONINE	MET	M
PHENYLALANINE	PHE	F
PROLINE	PRO	P
SERINE	SER	S
THREONINE	THR	T
TRYPTOPHAN	TRP	W
TYROSINE	TYR	Y
VALINE	VAL	V

ANTAGONIST shall mean materials (*e.g.*, ligands, candidate compounds) that competitively bind to the receptor at the same site as the agonists but which do not activate the intracellular response initiated by the active form of the receptor, and can thereby inhibit the intracellular responses by agonists. ANTAGONISTS do not diminish the baseline intracellular response in the absence of an agonist. In some embodiments, ANTAGONISTS are those materials not previously known to activate the intracellular response when they bind to the receptor or to enhance GTP binding to membranes.

CANDIDATE COMPOUND shall mean a molecule (for example, and not limitation, a chemical compound) that is amenable to a screening technique. Preferably, the phrase "candidate compound" does not include compounds which were publicly known to be compounds selected from the group consisting of inverse agonist, agonist or antagonist to a receptor, as previously determined by an indirect identification process ("indirectly identified compound"); more preferably, not including an indirectly identified compound which has previously been determined to have therapeutic efficacy in at least one mammal; and, most preferably, not including an indirectly identified compound which has previously been determined to have therapeutic utility in humans.

COMPOSITION means a material comprising at least one component; a “pharmaceutical composition” is an example of a composition.

COMPOUND EFFICACY shall mean a measurement of the ability of a compound to inhibit or stimulate receptor functionality; i.e. the ability to activate/inhibit a signal transduction pathway, as opposed to receptor binding affinity. Exemplary means of
5 detecting compound efficacy are disclosed in the Example section of this patent document.

CODON shall mean a grouping of three nucleotides (or equivalents to nucleotides) which generally comprise a nucleoside (adenosine (A), guanosine (G), cytidine (C), uridine (U) and thymidine (T)) coupled to a phosphate group and which, when translated, encodes
10 an amino acid.

CONSTITUTIVELY ACTIVATED RECEPTOR shall mean a receptor subjected to constitutive receptor activation. A constitutively activated receptor can be endogenous or non-endogenous.

CONSTITUTIVE RECEPTOR ACTIVATION shall mean stabilization of a
15 receptor in the active state by means other than binding of the receptor with its ligand or a chemical equivalent thereof.

CONTACT or **CONTACTING** shall mean bringing at least two moieties together, whether in an in vitro system or an in vivo system.

DIRECTLY IDENTIFYING or **DIRECTLY IDENTIFIED**, in relationship to
20 the phrase “candidate compound”, shall mean the screening of a candidate compound against a constitutively activated receptor, preferably a constitutively activated orphan receptor, and most preferably against a constitutively activated G protein-coupled cell surface orphan receptor, and assessing the compound efficacy of such compound. This

phrase is, under no circumstances, to be interpreted or understood to be encompassed by or to encompass the phrase "indirectly identifying" or "indirectly identified."

ENDOGENOUS shall mean a material that a mammal naturally produces. **ENDOGENOUS** in reference to, for example and not limitation, the term "receptor," shall mean that which is naturally produced by a mammal (for example, and not limitation, a human) or a virus. By contrast, the term **NON-ENDOGENOUS** in this context shall mean that which is not naturally produced by a mammal (for example, and not limitation, a human) or a virus. For example, and not limitation, a receptor which is not constitutively active in its endogenous form, but when manipulated becomes constitutively active, is most preferably referred to herein as a "non-endogenous, constitutively activated receptor." Both terms can be utilized to describe both "in vivo" and "in vitro" systems. For example, and not limitation, in a screening approach, the endogenous or non-endogenous receptor may be in reference to an in vitro screening system. As a further example and not limitation, where the genome of a mammal has been manipulated to include a non-endogenous constitutively activated receptor, screening of a candidate compound by means of an in vivo system is viable.

G PROTEIN COUPLED RECEPTOR FUSION PROTEIN and **GPCR FUSION PROTEIN**, in the context of the invention disclosed herein, each mean a non-endogenous protein comprising an endogenous, constitutively activate GPCR or a non-endogenous, constitutively activated GPCR fused to at least one G protein, most preferably the alpha (α) subunit of such G protein (this being the subunit that binds GTP), with the G protein preferably being of the same type as the G protein that naturally couples with endogenous orphan GPCR. For example, and not limitation, in an endogenous state, if the G protein " $G_s\alpha$ " is the predominate G protein that couples with the GPCR, a GPCR Fusion

Protein based upon the specific GPCR would be a non-endogenous protein comprising the GPCR fused to $G_s\alpha$; in some circumstances, as will be set forth below, a non-predominant G protein can be fused to the GPCR. The G protein can be fused directly to the C-terminus of the constitutively active GPCR or there may be spacers between the two.

5 **HOST CELL** shall mean a cell capable of having a Plasmid and/or Vector incorporated therein. In the case of a prokaryotic Host Cell, a Plasmid is typically replicated as a autonomous molecule as the Host Cell replicates (generally, the Plasmid is thereafter isolated for introduction into a eukaryotic Host Cell); in the case of a eukaryotic Host Cell, a Plasmid is integrated into the cellular DNA of the Host Cell such that when the
10 eukaryotic Host Cell replicates, the Plasmid replicates. In some embodiments the Host Cell is eukaryotic, more preferably, mammalian, and most preferably selected from the group consisting of 293, 293T and COS-7 cells.

INDIRECTLY IDENTIFYING or **INDIRECTLY IDENTIFIED** means the traditional approach to the drug discovery process involving identification of an endogenous
15 ligand specific for an endogenous receptor, screening of candidate compounds against the receptor for determination of those which interfere and/or compete with the ligand-receptor interaction, and assessing the efficacy of the compound for affecting at least one second messenger pathway associated with the activated receptor.

INHIBIT or **INHIBITING**, in relationship to the term "response" shall mean that a
20 response is decreased or prevented in the presence of a compound as opposed to in the absence of the compound.

INVERSE AGONISTS shall mean materials (*e.g.*, ligand, candidate compound) which bind to either the endogenous form of the receptor or to the constitutively activated form of the receptor, and which inhibit the baseline intracellular response initiated by the

active form of the receptor below the normal base level of activity which is observed in the absence of agonists, or decrease GTP binding to membranes. Preferably, the baseline intracellular response is inhibited in the presence of the inverse agonist by at least 30%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%,
5 at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, and most preferably at least 99% as compared with the baseline response in the absence of the inverse agonist.

KNOWN RECEPTOR shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has been identified.

10 **LIGAND** shall mean a molecule specific for a naturally occurring receptor.

MUTANT or **MUTATION** in reference to an endogenous receptor's nucleic acid and/or amino acid sequence shall mean a specified change or changes to such endogenous sequences such that a mutated form of an endogenous, non-constitutively activated receptor evidences constitutive activation of the receptor. In terms of equivalents to specific
15 sequences, a subsequent mutated form of a human receptor is considered to be equivalent to a first mutation of the human receptor if (a) the level of constitutive activation of the subsequent mutated form of a human receptor is substantially the same as that evidenced by the first mutation of the receptor; and (b) the percent sequence (amino acid and/or nucleic acid) homology between the subsequent mutated form of the receptor and the first mutation
20 of the receptor is at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, and most preferably at least 99%. In some embodiments, owing to the fact that some preferred cassettes disclosed herein for achieving constitutive activation include a single amino acid and/or codon change between the

endogenous and the non-endogenous forms of the GPCR, it is preferred that the percent sequence homology should be at least 98%.

NON-ORPHAN RECEPTOR shall mean an endogenous naturally occurring molecule specific for an identified ligand wherein the binding of a ligand to a receptor
5 activates an intracellular signaling pathway.

ORPHAN RECEPTOR shall mean an endogenous receptor for which the ligand specific for that receptor has not been identified or is not known.

PHARMACEUTICAL COMPOSITION shall mean a composition comprising at least one active ingredient, whereby the composition is amenable to investigation for a
10 specified, efficacious outcome in a mammal (for example, and not limitation, a human). Those of ordinary skill in the art will understand and appreciate the techniques appropriate for determining whether an active ingredient has a desired efficacious outcome based upon the needs of the artisan.

PLASMID shall mean the combination of a Vector and cDNA. Generally, a
15 Plasmid is introduced into a Host Cell for the purposes of replication and/or expression of the cDNA as a protein.

SECOND MESSENGER shall mean an intracellular response produced as a result of receptor activation. A second messenger can include, for example, inositol triphosphate (IP₃), diacylglycerol (DAG), cyclic AMP (cAMP), and cyclic GMP (cGMP). Second
20 messenger response can be measured for a determination of receptor activation. In addition, second messenger response can be measured for the direct identification of candidate compounds, including for example, inverse agonists, agonists, and antagonists.

SIGNAL TO NOISE RATIO shall mean the signal generated in response to activation, amplification, or stimulation wherein the signal is above the background noise or the basal level in response to non-activation, non-amplification, or non-stimulation.

SPACER shall mean a translated number of amino acids that are located after the
5 last codon or last amino acid of a gene, for example a GPCR of interest, but before the start codon or beginning regions of the G protein of interest, wherein the translated number amino acids are placed in-frame with the beginnings regions of the G protein of interest. The number of translated amino acids can be tailored according to the needs of the skilled artisan and is generally from about one amino acid, preferably two amino acids, more
10 preferably three amino acids, more preferably four amino acids, more preferably five amino acids, more preferably six amino acids, more preferably seven amino acids, more preferably eight amino acids, more preferably nine amino acids, more preferably ten amino acids, more preferably eleven amino acids, and even more preferably twelve amino acids.

STIMULATE or **STIMULATING**, in relationship to the term "response" shall
15 mean that a response is increased in the presence of a compound as opposed to in the absence of the compound.

SUBSTANTIALLY shall refer to a result which is within 40% of a control result, preferably within 35%, more preferably within 30%, more preferably within 25%, more preferably within 20%, more preferably within 15%, more preferably within 10%, more
20 preferably within 5%, more preferably within 2%, and most preferably within 1% of a control result. For example, in the context of receptor functionality, a test receptor may exhibit substantially similar results to a control receptor if the transduced signal, measured using a method taught herein or similar method known to the art-skilled, if within 40% of the signal produced by a control signal.

VECTOR in reference to cDNA shall mean a circular DNA capable of incorporating at least one cDNA and capable of incorporation into a Host Cell.

The order of the following sections is set forth for presentational efficiency and is not intended, nor should be construed, as a limitation on the disclosure or the claims to follow.

A. Introduction

The traditional study of receptors has typically proceeded from the *a priori* assumption (historically based) that the endogenous ligand must first be identified before discovery could proceed to find antagonists and other molecules that could affect the receptor. Even in cases where an antagonist might have been known first, the search immediately extended to looking for the endogenous ligand. This mode of thinking has persisted in receptor research even after the discovery of constitutively activated receptors. What has not been heretofore recognized is that it is the active state of the receptor that is most useful for discovering agonists and inverse agonists of the receptor. For those diseases which result from an overly active receptor or an under-active receptor, what is desired in a therapeutic drug is a compound which acts to diminish the active state of a receptor or enhance the activity of the receptor, respectively, not necessarily a drug which is an antagonist to the endogenous ligand. This is because a compound that reduces or enhances the activity of the active receptor state need not bind at the same site as the endogenous ligand. Thus, as taught by a method of this invention, any search for therapeutic compounds should start by screening compounds against the ligand-independent active state.

B. Identification of Human GPCRs

The efforts of the Human Genome project have led to the identification of a plethora of information regarding nucleic acid sequences located within the human genome; it has been the case in this endeavor that genetic sequence information has been made available without an understanding or recognition as to whether or not any particular genomic
5 sequence does or may contain open-reading frame information that translate human proteins. Several methods of identifying nucleic acid sequences within the human genome are within the purview of those having ordinary skill in the art.

Receptor homology is useful in terms of gaining an appreciation of a role of the receptors within the human body. As the patent document progresses, techniques for
10 mutating these receptors to establish non-endogenous, constitutively activated versions of these receptors will be discussed.

The techniques disclosed herein are also applicable to other human GPCRs known to the art, as will be apparent to those skilled in the art.

C. Receptor Screening

15 Screening candidate compounds against a non-endogenous, constitutively activated version of the GPCRs disclosed herein allows for the direct identification of candidate compounds which act at the cell surface receptor, without requiring use of the receptor's endogenous ligand. Using routine, and often commercially available techniques, one can determine areas within the body where the endogenous version of human GPCRs disclosed
20 herein is expressed and/or over-expressed. The expression location of a receptor in a specific tissue provides a scientist with the ability to assign a physiological functional role of the receptor. It is also possible using these techniques to determine related disease/disorder states which are associated with the expression and/or over-expression of the receptor; such an approach is disclosed in this patent document. Furthermore,

expression of a receptor in diseased organs can assist one in determining the magnitude of the clinical relevance of the receptor.

Constitutive activation of the GPCRs disclosed herein is based upon the distance from the proline residue at which is presumed to be located within TM6 of the GPCR; this
5 algorithmic technique is disclosed in co-pending and commonly assigned patent document PCT Application Number PCT/US99/23938, published as WO 00/22129 on April 20, 2000, which, along with the other patent documents listed herein, is incorporated herein by reference in its entirety. The algorithmic technique is not predicated upon traditional sequence "alignment" but rather a specified distance from the aforementioned TM6 proline
10 residue (or, of course, endogenous constitutive substitution for such proline residue). By mutating an amino acid of residue located 16 amino acid residues from this residue (presumably located in the IC3 region of the receptor) to, most preferably, a lysine residue, constitutive activation of the receptor may be obtained. Other amino acid residues may be useful in the mutation at this position to achieve this objective.

15 **D. Disease/Disorder Identification and/or Selection**

As will be set forth in greater detail below, inverse agonists and agonists to the non-endogenous, constitutively activated GPCR can be identified by the methodologies of this invention. Such inverse agonists and agonists are ideal candidates as lead compounds in drug discovery programs for treating diseases related to this receptor. Because of the ability
20 to directly identify inverse agonists to the GPCR, thereby allowing for the development of pharmaceutical compositions, a search for diseases and disorders associated with the GPCR is relevant. The expression location of a receptor in a specific tissue provides a scientist with the ability to assign a physiological function to the receptor. For example, scanning both diseased and normal tissue samples for the presence of the GPCR now becomes more

than an academic exercise or one which might be pursued along the path of identifying an endogenous ligand to the specific GPCR. Tissue scans can be conducted across a broad range of healthy and diseased tissues. Such tissue scans provide a potential first step in associating a specific receptor with a disease and/or disorder. Furthermore, expression of a
5 receptor in diseased organs can assist one in determining the magnitude of clinical relevance of the receptor. Skilled artisans, armed with the present specification, are credited with the ability to infer the function of a GPCR once the receptor is localized to a certain tissue or region.

The DNA sequence of the GPCR can be used to make a probe/primer. In some
10 preferred embodiments the DNA sequence is used to make a probe for (a) dot-blot analysis against tissue-mRNA, and/or (b) RT-PCR identification of the expression of the receptor in tissue samples. The presence of a receptor in a tissue source, or a diseased tissue, or the presence of the receptor at elevated concentrations in diseased tissue compared to a normal tissue, can be used to correlate location to function and indicate the receptor's physiological
15 role/function and create a treatment regimen, including but not limited to, a disease associated with that function/role. Receptors can also be localized to regions of organs by this technique. Based on the known or assumed roles/functions of the specific tissues to which the receptor is localized, the putative physiological function of the receptor can be deduced. For example and not limitation, proteins located/expressed in areas of the
20 thalamus are associated with sensorimotor processing and arousal (*see*, Goodman & Gilman's, The Pharmacological Basis of Therapeutics, 9th Edition, page 465 (1996)). Proteins expressed in the hippocampus or in Schwann cells are associated with learning and memory, and myelination of peripheral nerves, respectively (*see*, Kandel, E. et al., Essentials of Neural Science and Behavior pages 657, 680 and 28, respectively (1995)). In

some embodiments, the probes and/or primers may be used to detect and/or diagnose diseases and/or disorders, including but not limited to, those diseases and disorders identified in Example 6, *infra*. Methods of generating such primers and/or probes are well known to those of skill in the art as well as methods of using primers and/or probes to detect
5 diseases and/or disorders.

E. Screening of Candidate Compounds

1. Generic GPCR screening assay techniques

When a G protein receptor becomes constitutively active, it binds to a G protein
10 (*e.g.*, G_q, G_s, G_i, G_z, G_o) and stimulates the binding of GTP to the G protein. The G protein then acts as a GTPase and hydrolyzes the GTP to GDP, whereby the receptor, under normal conditions, becomes deactivated. However, constitutively activated receptors continue to exchange GDP to GTP. A non-hydrolyzable analog of GTP, [³⁵S]GTPγS, can be used to monitor enhanced binding to membranes which express constitutively activated receptors.
15 It is reported that [³⁵S]GTPγS can be used to monitor G protein coupling to membranes in the absence and presence of ligand. An example of this monitoring, among other examples well-known and available to those in the art, was reported by Traynor and Nahorski in 1995. The use of this assay system is typically for initial screening of candidate compounds because the system is generically applicable to all G protein-coupled receptors regardless of
20 the particular G protein that interacts with the intracellular domain of the receptor.

2. Specific GPCR screening assay techniques

Once candidate compounds are identified using the “generic” G protein-coupled receptor assay (*i.e.*, an assay to select compounds that are agonists or inverse agonists), further screening to confirm that the compounds have interacted at the receptor site is

preferred. For example, a compound identified by the "generic" assay may not bind to the receptor, but may instead merely "uncouple" the G protein from the intracellular domain.

a. G_s, G_z and G_i.

G_s stimulates the enzyme adenylyl cyclase. G_i (and G_z and G_o), on the other hand,
5 inhibits adenylyl cyclase. Adenylyl cyclase catalyzes the conversion of ATP to cAMP;
thus, constitutively activated GPCRs that couple the G_s protein are associated with
increased cellular levels of cAMP. On the other hand, constitutively activated GPCRs that
couple G_i (or G_z, G_o) protein are associated with decreased cellular levels of cAMP. *See,*
10 *generally*, "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To
Brain (3rd Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992). Thus, assays that
detect cAMP can be utilized to determine if a candidate compound is, *e.g.*, an inverse
agonist to the receptor (*i.e.*, such a compound would decrease the levels of cAMP). A
variety of approaches known in the art for measuring cAMP can be utilized; a most
preferred approach relies upon the use of anti-cAMP antibodies in an ELISA-based format.
15 Another type of assay that can be utilized is a whole cell second messenger reporter system
assay. Promoters on genes drive the expression of the proteins that a particular gene
encodes. Cyclic AMP drives gene expression by promoting the binding of a cAMP-
responsive DNA binding protein or transcription factor (CREB) that then binds to the
promoter at specific sites (cAMP response elements) and drives the expression of the gene.
20 Reporter systems can be constructed which have a promoter containing multiple cAMP
response elements before the reporter gene, *e.g.*, β -galactosidase or luciferase. Thus, a
constitutively activated G_s-linked receptor causes the accumulation of cAMP that then
activates the gene and leads to the expression of the reporter protein. The reporter protein

such as β -galactosidase or luciferase can then be detected using standard biochemical assays (Chen et al. 1995).

b. G_o and G_q

5 G_q and G_o are associated with activation of the enzyme phospholipase C, which in turn hydrolyzes the phospholipid PIP₂, releasing two intracellular messengers: diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP₃). Increased accumulation of IP₃ is associated with activation of G_q- and G_o-associated receptors. *See, generally*, "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To Brain (3rd Ed.)
10 Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992). Assays that detect IP₃ accumulation can be utilized to determine if a candidate compound is, *e.g.*, an inverse agonist to a G_q- or G_o-associated receptor (*i.e.*, such a compound would decrease the levels of IP₃). G_q-associated receptors can also be examined using an AP1 reporter assay wherein G_q-dependent phospholipase C causes activation of genes containing AP1 elements; thus,
15 activated G_q-associated receptors will evidence an increase in the expression of such genes, whereby inverse agonists thereto will evidence a decrease in such expression, and agonists will evidence an increase in such expression. Commercially available assays for such detection are available.

3. GPCR Fusion Protein

20 The use of an endogenous, constitutively activated GPCR or a non-endogenous, constitutively activated GPCR, for use in screening of candidate compounds for the direct identification of inverse agonists, agonists provide an interesting screening challenge in that, by definition, the receptor is active even in the absence of an endogenous ligand bound thereto. Thus, in order to differentiate between, *e.g.*, the non-endogenous receptor in the
25 presence of a candidate compound and the non-endogenous receptor in the absence of that

compound, with an aim of such a differentiation to allow for an understanding as to whether such compound may be an inverse agonist or agonist or have no affect on such a receptor, it is preferred that an approach be utilized that can enhance such differentiation. A preferred approach is the use of a GPCR Fusion Protein.

5 Generally, once it is determined that a non-endogenous GPCR has been constitutively activated using the assay techniques set forth above (as well as others), it is possible to determine the predominant G protein that couples with the endogenous GPCR. Coupling of the G protein to the GPCR provides a signaling pathway that can be assessed. In some embodiments it is preferred that screening take place using a mammalian
10 expression system, such a system will be expected to have endogenous G protein therein. Thus, by definition, in such a system, the non-endogenous, constitutively activated GPCR will continuously signal. In some embodiments it is preferred that this signal be enhanced such that in the presence of, *e.g.*, an inverse agonist to the receptor, it is more likely that it will be able to more readily differentiate, particularly in the context of screening, between
15 the receptor when it is contacted with the inverse agonist.

 The GPCR Fusion Protein is intended to enhance the efficacy of G protein coupling with the non-endogenous GPCR. The GPCR Fusion Protein is preferred for screening with either an endogenous, constitutively active GPCR or a non-endogenous, constitutively activated GPCR because such an approach increases the signal that is utilized in such
20 screening techniques. This is important in facilitating a significant "signal to noise" ratio; such a significant ratio is preferred for the screening of candidate compounds as disclosed herein.

 The construction of a construct useful for expression of a GPCR Fusion Protein is within the purview of those having ordinary skill in the art. Commercially available

expression vectors and systems offer a variety of approaches that can fit the particular needs of an investigator. Important criteria on the construction of such a GPCR Fusion Protein construct include but are not limited to, that the endogenous GPCR sequence and the G protein sequence both be in-frame (preferably, the sequence for the endogenous GPCR is upstream of the G protein sequence), and that the "stop" codon of the GPCR be deleted or replaced such that upon expression of the GPCR, the G protein can also be expressed. Other embodiments include constructs wherein the endogenous GPCR sequence and the G protein sequence are not in-frame and/or the "stop" codon is not deleted or replaced. The GPCR can be linked directly to the G protein, or there can be spacer residues between the two (preferably, no more than about 12, although this number can be readily ascertained by one of ordinary skill in the art). Based upon convenience it is preferred to use a spacer. Preferably, the G protein that couples to the non-endogenous GPCR will have been identified prior to the creation of the GPCR Fusion Protein construct. Because there are only a few G proteins that have been identified, it is preferred that a construct comprising the sequence of the G protein (*i.e.*, a universal G protein construct (*see Examples*)) be available for insertion of an endogenous GPCR sequence therein; this provides for further efficiency in the context of large-scale screening of a variety of different endogenous GPCRs having different sequences.

As noted above, constitutively activated GPCRs that couple to G_i , G_z and G_o are expected to inhibit the formation of cAMP making assays based upon these types of GPCRs challenging (*i.e.*, the cAMP signal decreases upon activation thus making the direct identification of, *e.g.*, inverse agonists (which would further decrease this signal), challenging. As will be disclosed herein, we have ascertained that for these types of receptors, it is possible to create a GPCR Fusion Protein that is not based upon the GPCRs

endogenous G protein, in an effort to establish a viable cyclase-based assay. Thus, for example, an endogenous G_i coupled receptor can be fused to a G_s protein –such a fusion construct, upon expression, “drives” or “forces” the endogenous GPCR to couple with, *e.g.*, G_s rather than the “natural” G_i protein, such that a cyclase-based assay can be established.

- 5 Thus, for G_i , G_z and G_o coupled receptors, in some embodiments it is preferred that when a GPCR Fusion Protein is used and the assay is based upon detection of adenylyl cyclase activity, that the fusion construct be established with G_s (or an equivalent G protein that stimulates the formation of the enzyme adenylyl cyclase).

G protein	Effect of cAMP Production upon Activation of GPCR (<i>i.e.</i> , constitutive activation or agonist binding)	Effect of IP_3 Accumulation upon Activation of GPCR (<i>i.e.</i> , constitutive activation or agonist binding)	Effect of cAMP Production upon contact with an Inverse Agonist	Effect on IP_3 Accumulation upon contact with an Inverse Agonist
G_s	Increase	N/A	Decrease	N/A
G_i	Decrease	N/A	Increase	N/A
G_z	Decrease	N/A	Increase	N/A
G_o	Decrease	Increase	Increase	Decrease
G_q	N/A	Increase	N/A	Decrease

10

- Equally effective is a G Protein Fusion construct that utilizes a G_q Protein fused with a G_s , G_i , G_z or G_o Protein. In some embodiments a preferred fusion construct can be accomplished with a G_q Protein wherein the first six (6) amino acids of the G-protein α -subunit (“ $G\alpha_q$ ”) is deleted and the last five (5) amino acids at the C-terminal end of $G\alpha_q$ is
- 15 replaced with the corresponding amino acids of the $G\alpha$ of the G protein of interest. For example, a fusion construct can have a G_q (6 amino acid deletion) fused with a G_i Protein, resulting in a “ G_q/G_i Fusion Construct”. This fusion construct will forces the endogenous G_i coupled receptor to couple to its non-endogenous G protein, G_q , such that the second

messenger, for example, inositol triphosphate or diacylglycerol, can be measured *in lieu* of cAMP production.

4. Co-transfection of a Target G_i Coupled GPCR with a Signal-Enhancer G_s Coupled GPCR (cAMP Based Assays)

5

A G_i coupled receptor is known to inhibit adenylyl cyclase, and, therefore, decreases the level of cAMP production, which can make assessment of cAMP levels challenging. An effective technique in measuring the decrease in production of cAMP as an indication of constitutive activation of a receptor that predominantly couples G_i upon activation can be accomplished by co-transfecting a signal enhancer, *e.g.*, a non-endogenous, constitutively activated receptor that predominantly couples with G_s upon activation (*e.g.*, TSHR-A623I, disclosed below), with the G_i linked GPCR. As is apparent, constitutive activation of a G_s coupled receptor can be determined based upon an increase in production of cAMP. Constitutive activation of a G_i coupled receptor leads to a decrease in production cAMP. Thus, the co-transfection approach is intended to advantageously exploit these “opposite” affects. For example, co-transfection of a non-endogenous, constitutively activated G_s coupled receptor (the “signal enhancer”) with the endogenous G_i coupled receptor (the “target receptor”) provides a baseline cAMP signal (*i.e.*, although the G_i coupled receptor will decrease cAMP levels, this “decrease” will be relative to the substantial increase in cAMP levels established by constitutively activated G_s coupled signal enhancer). By then co-transfecting the signal enhancer with a constitutively activated version of the target receptor, cAMP would be expected to further decrease (relative to base line) due to the increased functional activity of the G_i target (*i.e.*, which decreases cAMP).

Screening of candidate compounds using a cAMP based assay can then be accomplished, with two 'changes' relative to the use of the endogenous receptor/G-protein fusion: first, relative to the G_i coupled target receptor, “opposite” effects will result, *i.e.*, an

25

inverse agonist of the G_i coupled target receptor will increase the measured cAMP signal, while an agonist of the G_i coupled target receptor will decrease this signal; second, as would be apparent, candidate compounds that are directly identified using this approach should be assessed independently to ensure that these do not target the signal enhancing receptor (this
5 can be done prior to or after screening against the co-transfected receptors).

F. Medicinal Chemistry

Generally, but not always, direct identification of candidate compounds is conducted in conjunction with compounds generated via combinatorial chemistry
10 techniques, whereby thousands of compounds are randomly prepared for such analysis. Generally, the results of such screening will be compounds having unique core structures; thereafter, these compounds may be subjected to additional chemical modification around a preferred core structure(s) to further enhance the medicinal properties thereof. Such techniques are known to those in the art and will not be
15 addressed in detail in this patent document.

G. Pharmaceutical compositions

Candidate compounds selected for further development can be formulated into pharmaceutical compositions using techniques well known to those in the art. Suitable
20 pharmaceutically-acceptable carriers are available to those in the art; for example, see Remington's Pharmaceutical Sciences, 16th Edition, 1980, Mack Publishing Co., (Osol et al., eds.).

H. Other Utilities

Although a preferred use of the non-endogenous versions of the GPCRs disclosed herein may be for the direct identification of candidate compounds as inverse agonists or agonists (preferably for use as pharmaceutical agents), other uses of these versions of GPCRs exist. For example, *in vitro* and *in vivo* systems incorporating GPCRs can be
5 utilized to further elucidate and understand the roles these receptors play in the human condition, both normal and diseased, as well as understanding the role of constitutive activation as it applies to understanding the signaling cascade. In some embodiments it is preferred that the endogenous receptors be "orphan receptors", *i.e.*, the endogenous ligand for the receptor has not been identified. In some embodiments, therefore, the modified,
10 non-endogenous GPCRs can be used to understand the role of endogenous receptors in the human body before the endogenous ligand therefore is identified. Such receptors can also be used to further elucidate known receptors and the pathways through which they transduce a signal. Other uses of the disclosed receptors will become apparent to those in the art based upon, *inter alia*, a review of this patent document.

15

EXAMPLES

The following examples are presented for purposes of elucidation, and not limitation, of the present invention. While specific nucleic acid and amino acid sequences are disclosed herein, those of ordinary skill in the art are credited with the ability to make
20 minor modifications to these sequences while achieving the same or substantially similar results reported below. The traditional approach to application or understanding of sequence cassettes from one sequence to another (*e.g.* from rat receptor to human receptor or from human receptor A to human receptor B) is generally predicated upon sequence alignment techniques whereby the sequences are aligned in an effort to determine areas of

commonality. The mutational approach disclosed herein does not rely upon this approach but is instead based upon an algorithmic approach and a positional distance from a conserved proline residue located within the TM6 region of human GPCRs. Once this approach is secured, those in the art are credited with the ability to make minor
 5 modifications thereto to achieve substantially the same results (*i.e.*, constitutive activation) disclosed herein. Such modified approaches are considered within the purview of this disclosure.

Example 1

ENDOGENOUS HUMAN GPCRS

10 The following cDNA receptors were cloned by utilizing the techniques in this Section, see below. Table B lists the receptors disclosed throughout this patent applications, the open reading frame, the nucleic acid and the amino acid sequences for the endogenous GPCR (Table B).

TABLE B

Disclosed Human GPCRs	Open Reading Frame (Base Pairs)	Nucleic Acid SEQ.ID. NO.	Amino Acid SEQ.ID.NO.
FPRL-2	1,062bp	1	2
STLR33	1,029bp	3	4
GPR45	1,119bp	5	6
mGluR7	2,748bp	7	8
GPR37	1,842bp	9	10
HF1948	1,086bp	11	12
GPR66	957bp	13	14
GPR35	930bp	15	16
ETBR-LP2	1,446bp	17	18
GPR26	1,011	97	98

15

2. Full Length Cloning Protocol

a. FPRL-2 (Seq. Id. Nos. 1 & 2)

FPRL-2 was cloned and sequenced in 1992. Bao, L. et al, 13(2) *Genomics* 437-40 (1992). FPRL-2 has been reported to be located on chromosome 19 having a sequence similarity to N-formyl peptide receptor like-1 (FPRL-1) both of which share significant similarity with the N-formyl peptide receptor (FPR). The endogenous ligand for FPR is formyl peptide, however, the two homologues of FPR, FPRL-1 and FPRL-2, do not bind to the same ligand but are likely chemotactic receptors. 13(2) *Genomics* 437-40 (1992). Chemotactic receptors are reported to be involved in inflammation. FPRL-2 is a GPCR having an open reading frame of 1062 bp encoding a 353 amino acid protein.

10 PCR was performed using genomic cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μ M of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 64°C for 1 min 20 sec and 72°C for 2 min. The 5' PCR contained an EcoRI site with the following sequence

15 5'-AAAGATTCAGGTGTGGGAAGATGGAAACC-3' (SEQ.ID.NO.:19)
and the 3' primer contained an ApaI site with the following sequence:
5'-AAAGGATCCCCGACCTCACATTGCTTGTA -3' (SEQ.ID.NO.:20).

The PCR fragment was digested with EcoRI and ApaI and cloned into an EcoRI-ApaI site of CMV expression vector. Nucleic acid (SEQ.ID.NO.:1) and amino acid
20 (SEQ.ID.NO.:2) sequences for human FPRL-2 were thereafter determined and verified.

b. STLR33 (Seq. Id. Nos. 3 & 4)

PCR was performed using genomic cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μ M of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C

for 1 min, 62°C for 1 min 20 sec and 72°C for 2 min. The 5' PCR contained an EcoRI site with the following sequence

5'-CAGGAATTCATCAGAACAGACACCATGGCA-3' (SEQ.ID.NO.:21)

and the 3' primer contained a BamHI site with the following sequence:

5 5'-GCAGGATCCAGAGCAGTTTTTTTCGAAACCCT -3' (SEQ.ID.NO.:22).

The PCR fragment was digested with EcoRI and BamHI and cloned into an EcoRI-BamHI site of CMV expression vector. Nucleic acid (SEQ.ID.NO.:3) and amino acid (SEQ.ID.NO.:4) sequences for human STRL33 were thereafter determined and verified.

10 c. GPR45 (Seq. Id. Nos. 5 & 6)

PCR was performed using genomic cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 µM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was as follows with cycles 2 through four repeated 35 times: 96°C for 2 min, 96°C for 30 sec, 55°C for 20
15 sec. 72°C for 1 min and 20 sec, and 72°C for 5 min. The 5' PCR contained a HindIII site with the following sequence

5'-TCCAAGCTTCAAGGGTCTCTCCACGATGGCCTG-3' (SEQ.ID.NO.:23)

and the 3' primer contained an EcoRI site with the following sequence:

5'-TGCGAATTCTCTGTGGCCCCCTGACCCCTAAA -3' (SEQ.ID.NO.:24).

20 The PCR fragment was digested with HindIII and EcoRI and cloned into a HindIII-EcoRI site of CMV expression vector. Nucleic acid (SEQ.ID.NO.:5) and amino acid (SEQ.ID.NO.:6) sequences for human GPR45 were thereafter determined and verified.

The cDNA was then tagged with V5 by resubcloning into V5-His vector using pfu PCR and the following two primers utilized had the following sequence:

5'-GGTAAGCTTACCATGGCCTGCAACAGCACGTCCCTT-3' (SEQ.ID.NO.:25) and

5'-GACGAATTCAACCGCAGACTGGTITTCATTGCA-3' (SEQ.ID.NO.:26).

5 The cycle condition was 30 cycles of 94°C for 1 min, 60°C for 2min and 72°C for 2 min.

d. mGLUR7 (Seq. Id. Nos. 7 & 8)

Glutamate is an excitatory neurotransmitter which is abundantly found in the mammalian brain. *See*, Dingledine, R. et al., 130(4S Suppl) J Nutr. 1039S (2000). There
10 are two classes of glutamate receptor, the ionotropic (ligand-gated ion channels) and the metabotropic (GPCRs). Metabotropic glutamate receptors are a heterogenous family of GPCRs that are linked to several second messenger pathways to regulate neuronal excitability and synaptic transmission. (*See*, Phillips, T. et al., 57(1) Brain Res Mol Brain Res 132 (1998)). Metabotropic glutamate receptor type 7 (mGluR7) has been reported to be
15 expressed in the brain, with highest levels of expression found in the hippocampus, cerebral cortex and cerebellum. *See*, Makoff, A. et al., 40(1) Brain Res Mol Brain Res 165 (1996). Based on the areas of the brain in which the receptor is localized, the putative functional role of the receptor can be deduced. For example, and while not wishing to be bound by any particular theory, mGluR7 is thought to play a role in depression, anxiety, obesity,
20 Alzheimer's Disease, pain and stroke.

mGluR7 cDNA was generously supplied by Elizabeth Hoffman, Ph.D. The vector utilized for mGluR7 was pRcCMV (the coding region for mGluR7 was subcloned into pCMV vector at an EcoRI-ClaI site). *See*, SEQ.ID.NO.:7 for nucleic acid sequence and SEQ.ID.NO.:8 for the deduced amino acid sequence of mGluR7.

e. GPR37 (Seq. Id. Nos. 9 & 10)

The present invention also relates to the human GPR37. GPR37 was cloned and sequenced in 1997. Marazziti, D. et al, 45 (1) *Genomics* 68-77 (1997). GPR37 is an orphan GPCR having an open reading frame of 1839 bp encoding a 613 amino acid protein.

5 GPR37 has been reported to share homology with the endothelin type B-like receptor and expressed in the human brain tissues, particularly in corpus callosum, medulla, putamen, and caudate nucleus.

PCR was performed using brain cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μ M of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 62°C for 1min and 72°C for 2 min. The 5' PCR contained a HindIII site with the following sequence

5'-GCAAGCTTGTGCCCTCACCAAGCCATGCGAGCC-3' (SEQ.ID.NO.:27)

and the 3' primer contained an EcoRI site with the following sequence:

15 5'-CGGAATTCAGCAATGAGTTCCGACAGAAGC -3' (SEQ.ID.NO.:28).

The 1.9 kb PCR fragment was digested with HindIII and EcoRI and cloned into a HindIII-EcoRI site of CMVp expression vector. Nucleic acid (SEQ.ID.NO.:9) and amino acid (SEQ.ID.NO.:10) sequences for human GPR37 were thereafter determined and verified.

20 f. HF1948 (Seq. Id. Nos. 11 & 12)

HF1948 cDNA was generously supplied by Elizabeth Hoffman, Ph.D. The vector utilized for HF1948 was pRcCMV (the coding region for HF1948 was subcloned into pCMV vector at an HindIII-BamHI site). See, SEQ.ID.NO.:11 for nucleic acid sequence and SEQ.ID.NO.:12 for the deduced amino acid sequence of HF1948.

g. GPR66 (Seq. Id. Nos. 13 & 14)

The cDNA for human GPR66 (GenBank Accession Numbers AF044600 and AF044601) was generated and cloned into pCMV expression vector as follows: PCR was performed using genomic DNA as template and TaqPlus Precision polymerase (Stratagene)

5 for first round PCR or pfu polymerase (Stratagene) for second round PCR with the buffer system provided by the manufacturer, 0.25 μ M of each primer, and 0.2 mM (TaqPlus Precision) or 0.5 mM (pfu) of each of the 4 nucleotides. When pfu was used, 10% DMSO was included in the buffer. The cycle condition was 30 cycles of: 94°C for 1 min; 65°C for 1min; and 72°C for: (a) 1 min for first round PCR; and (b) 2 min for second round PCR.

10 Because there is an intron in the coding region, two sets of primers were separately used to
generate overlapping 5' and 3' fragments. The 5' fragment PCR primers were:

5'-ACCATGGCTTGCAATGGCAGTGCGGCCAGGGGGCACT-3' (external sense)

(SEQ.ID.NO.:29) and

5'-CGACCAGGACAAACAGCATCTTGGTCACTTGTCTCCGGC-3' (internal antisense)

15 (SEQ.ID.NO.:30).

The 3' fragment PCR primers were:

5'-GACCAAGATGCTGTTTGTCCTGGTCGTGGTGTTTGGCAT-3' (internal sense)

(SEQ.ID.NO.:31) and

5'-CGGAATTCAGGATGGATCGGTCTCTTGCTGCGCCT-3' (external antisense with an EcoRI

20 site) (SEQ.ID.NO.:32).

The 5' and 3' fragments were ligated together by using the first round PCR as template and the kinased external sense primer and external antisense primer to perform second round PCR. The 1.2 kb PCR fragment was digested with EcoRI and cloned into the blunt-EcoRI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.:13) and amino

acid (SEQ.ID.NO.:14) sequences for human GPR66 were thereafter determined and verified.

h. GPR35 (Seq. Id. Nos. 15 & 16)

GPR35 is a 309 amino acid sequence whereby the endogenous ligand for GPR35 is unknown (O'Dowd B. et al., 47(2) Genomics 310 (1998)). GPR35 was determined to interact with a specific transcription factor, known as E2F, which is necessary for initiating DNA replication and, ultimately, cell proliferation. Within a cell, E2F couples to a tumor suppressor gene, known as retino-blastoma ("Rb"). Upon phosphorylation of this transcription factor construct, E2F is liberated from the Rb gene and then enters the nucleus of the cell. Inside the nucleus, E2F binds to genes, such as DNA polymerase, to initiate DNA replication, which results in proliferation of the cell.

PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μ M of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 62°C for 1min and 72 °C for 1 min and 20 sec. The 5' PCR primer was kinased with the following sequence:

5'-GCGAATTCCGGCTCCCTGTGCTGCCCCAGG-3' (SEQ.ID.NO.:33)

and the 3' primer contains a BamHI site with the following sequence:

5'-GCGGATCCCGGAGCCCCGAGACCTGGCCC -3' (SEQ.ID.NO.:34).

The 1 kb PCR fragment was digested with BamHI and cloned into EcoRV-BamHI site of CMVp expression vector. All 6 clones sequenced contain a potential polymorphism involving change of amino acid 294 from Arg to Ser. Nucleic acid (SEQ.ID.NO.:15) and amino acid (SEQ.ID.NO.:16) sequences for human GPR35 were thereafter determined and verified.

i. ETBR-LP2 (Seq. Id. Nos. 17 & 18)

ETBR-LP2 was cloned and sequenced in 1998. Valdenaire O. et al, 424(3) *FEBS Lett.* 193 (1998); see Figure 1 of Valdenaire for deduced nucleic and amino acid sequences. ETBR-LP2 has an open reading frame of 1839 bp encoding a 613 amino acid protein.

5 ETBR-LP2 has been reported to share homology with the endothelin type B receptor (ETBR-LP). Further, ETBR-LP2 evidences about a 47% amino acid sequence homology with human GPR37. ETBR-LP2 has been reported to be expressed in the human central nervous system (e.g., cerebral cortex, internal capsule fibers and Bergmann glia (424 *FEBS Lett* at 196).

10 PCR was performed using brain cDNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μ M of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 65°C for 1min and 72 °C for 1.5 min. The 5' PCR contained an EcoRI site with the sequence:

5'-CTGGAATTCTCTGCTCATCCAGCCATGCGG -3' (SEQ.ID.NO.:35)

15 and the 3' primer contained a BamHI site with the sequence:

5'-CCTGGATCCCCACCCCTACTGGGGCCTCAG -3' (SEQ.ID.NO.:36).

The resulting 1.5 kb PCR fragment was digested with EcoRI and BamHI and cloned into EcoRI-BamHI site of pCMV expression vector. Nucleic acid (SEQ.ID.NO.:17) and amino acid (SEQ.ID.NO.:18) sequences for human ETBR-LP2 were thereafter determined

20 and verified.

j. GPR26 (Seq. Id. Nos. 97 & 98)

EST clone HIBB055, a 3' 400bp PCR fragment used to screen the Human Genomic lambda Dash II Library (Stratagene catalog special order). The screening conditions were as follows: filters were hybridize overnight at 55°C in a formamide based hybridization

solution. The washing conditions were 2X SSC/1%SDS twice at 65° and .2X SSC/.1%SDS twice at 65°C for 20min at each wash. The filters were placed on film exposed overnight at -80°C and developed the next day. The positive plaques were further characterized by a second round of phage screening from the primary plugs under the same conditions.

- 5 Human Fetal Brain cDNA library Uni-ZAP XR Vector (catalog#937227, Stratagene) was then probed with a 250bp probe generated from new sequence from the genomic library screening. The 250bp probe was generated by PCR with *Taqplus Precision* PCR system (Stratagene #600210) with manufacturer supplied buffer system. The cycling parameters were as follows: 30 cycles with 95°C for 45sec, 55°C for 40sec, 72°C for 1min
10 and final extension for 10 min. The primers utilized were as follows:
5'-CGAGAAGGTGCTCAAGGTGGC-3' (SEQ.ID.NO.: 99) and
5'-GAGAAGAGCTCCACTAGCCTGGTGATCACA-3' (SEQ. ID.NO.:100).

- The Human Fetal Brain cDNA library was probed with the same 250bp PCR fragment under the same conditions as the genomic library except the hybridization temp
15 was 42°C. The positive primary plugs were further characterized by a second round of screening under the same conditions with a hybridization temp. of 55°C. Positive plaques were analyzed by sequence via Sanger method and the start codon was obtained from one of the positive clones

- The human GPR26 full length clone was then generated by PCR using PfuTurbo
20 DNA Polymerase (Stratagene #600250) with the following parameters:
40 cycles of 95°C for 45 sec., 62°C for 1 min. and 72°C for 1.2 min. and a final extension of 10 min. at 72°C. The template used was Human Fetal Brain cDNA (Clontech# 7402-1) and the primers were as follows:

5'-GAATTCATGAACTCGTGGGACGCGGGCCTGGCGGGC-3' (SEQ.ID.NO.:101)

and

5'-CTCGAGTCACTCAGACACCGGCAGAATGTTCT-3' (SEQ.ID.NO.:102).

The fragment generated had a 5' EcoR1 linker and a 3' Xho1 linker. The PCR
5 product was digested using the given linker enzymes and subcloned into the expression
vector pcDNA3.1(+) (Invitrogen#V790-20) at the EcoR1/Xho1 sites using the Rapid
Ligation Kit (Roche#1635 379). Nucleic acid (SEQ.ID.NO.:97) and amino acid
(SEQ.ID.NO.:98) sequences for human GPR26 were thereafter determined and verified.

Example 2

10 PREPARATION OF NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED GPCRS

Those skilled in the art are credited with the ability to select techniques for
mutation of a nucleic acid sequence. Presented below are approaches utilized to create
non-endogenous versions of several of the human GPCRs disclosed above. The
mutations disclosed below are based upon an algorithmic approach whereby the 16th
15 amino acid (located in the IC3 region of the GPCR) from a conserved proline (or an
endogenous, conservative substitution therefore) residue (located in the TM6 region of
the GPCR, near the TM6/IC3 interface) is mutated, preferably to an alanine, histamine,
arginine or lysine amino acid residue, most preferably to a lysine amino acid residue.

1. Site-Directed Mutagenesis

20 Preparation of non-endogenous human GPCRs was accomplished on human
GPCRs using, *inter alia*, Transformer Site-Directed™ Mutagenesis Kit (Clontech)
according to the manufacturer instructions or QuikChange™ Site-Directed™ Mutagenesis
Kit (Stratagene, according to manufacturer's instructions). The following GPCRs were
mutated according with the above method using the designated sequence primers (Table C).

For convenience, the codon mutation to be incorporated into the human GPCR is also noted, in standard form (Table C):

TABLE C

Receptor Identifier	Codon Mutation	5'-3' orientation, mutation sequence underlined (SEQ.ID.NO.)	5'-3' orientation (SEQ.ID.NO.)
FLPR-2	T240K	TCCAGCCGTC <u>CCAA</u> CGT GTCTTCGCTGC (37)	CTCCTTCGGTCCTCCTA TCGTTGTCAGAAGT (38)
STRL33	L230K	CAGAAGCACAGATCAAA AAAGATCATCTTCCTG (39)	CTCCTTCGGTCCTCCTA TCGTTGTCAGAAGT (38)
mGluR7	W590S	AGTGGCACTCCOCCTCG GCTGTGATTCTCTGT (59)	ACAGGAATCACAGCC GAGGGGGAGTGCCAC T (40)
	R659H	TGTGTTCTTTCCGGCATG TTTTCTTGGGCTTG (41)	CAAGCCCAAGAAAAC ATGCCGGAAGAACA CA (42)
	T771C	CTCATGGTCACATGTTGT GTGTATGCCATCAAG (43)	CTTGATGGCATAACACA CAACATGTGACCATGA G (44)
	I790K	ACGAAGCCAAGCCCAAG GGATTCACTATGTACAC (45)	GTGTACATAGTGAATC CCTTGGGCTTGGCTCC GT (46)
GPR37	L352R	GTCACCACCTTTCACCCG ATGTGCTCTGTGCATAG (47)	CTATGCACAGAGCAC ATCGGGTGAAAGGTG GTGAC (48)
	C543Y	CCTTTGTCTTTAAGTC CTATGTCACCCAGTCCT (49)	AGGACTGGGGTGACA TAGGACTTAAAGAAC AAAAGG (50)
HF1948	I281F	ATGTGGAGCCCCATCTT CATCACCATCCTCC (51)	GGAGGATGGTGATGA AGATGGGGCTCCACAT (52)
	E135N	GCCGCGGTGAGCCTGAA TCGCATGGTGTGCATC (53)	GATGCACACCATGCG ATCAGGCTGACCGCG GC (54)
GPR66	T273K	GGCCGGAGACAAGTGAA AAGATGCTGTTT (55)	AAACAGCATCTTTTC ACTTGTCTCCGGCC (56)
GPR35	A216K	<i>See alternate approaches</i>	<i>See alternate approaches</i>
ETBR-LP2	N358K	GAGAGCCAGCTCAAGAG CACCGTGGTG (57)	CTCCTTCGGTCCTCCTA TCGTTGTCAGAAGT (58)

1. Alternative Approaches For Creation of Non-Endogenous Human GPCRs

Preparation of the non-endogenous, constitutively activated human GPR35 receptor was accomplished by creating a A216K mutation. Mutagenesis was performed using Transformer Site-Directed™ Mutagenesis Kit (Clontech) according to manufacturer's instructions. (see, SEQ.ID.NO.:84 for nucleic acid sequence, SEQ.ID.NO.:85 for amino acid sequence). The two mutagenesis primers were utilized, a lysine mutagenesis oligonucleotide and a selection marker oligonucleotide, which had the following sequences:

5'- GCCACCCGCAAGGCTAAACGCATGGTCTGG -3' (SEQ.ID.NO.:60 sense) and
5'- CTCCTTCGGTCCTCCTATCGTTGTCAGAAGT -3' (SEQ.ID.NO.:61; antisense),
10 respectively.

For first round PCR, SEQ.ID.NO.:33 and SEQ.ID.NO.:61 were used to generate the 5' 700 bp fragment, while SEQ.ID.NO.:34 and SEQ.ID.NO.:60 were used to generate the 3' 350 bp fragment. PCR was performed using endogenous GPR35 cDNA as template and pfu polymerase (Stratagene) with the buffer system provided by the manufacturer supplemented with 10% DMSO, 0.25 μ M of each primer, and 0.5 mM of each 4
15 nucleotides. The cycle condition was 25 cycles of 94°C for 30 sec, 65°C for 1min and 72 °C for 2 min and 20 sec. The 5' and 3' PCR fragment from first round PCR were then used as cotemplate to perform second round PCR using oligo 1 and 2 as primers and pfu polymerase as described above except the annealing temperature was 55 °C, and the
20 extention time was 2 min. The resulting PCR fragment was then digested and subcloned into pCMV as described for the endogenous cDNA.

The non-endogenous human GPCRs were then sequenced and the derived and verified nucleic acid and amino acid sequences are listed in the accompanying "Sequence Listing" appendix to this patent document, as summarized in Table D below:

TABLE D

Non—Endogenous Receptor	Nucleic Acid Sequence Listing	Amino Acid Sequence Listing
FPRL-2 L240K	SEQ.ID.NO.:62	SEQ.ID.NO.:63
STRL33 L230K	SEQ.ID.NO.:64	SEQ.ID.NO.:65
MgluR7 W590S R659H T771C I790K	SEQ.ID.NO.: 66 SEQ.ID.NO.:68 SEQ.ID.NO.:70 SEQ.ID.NO.:72	SEQ.ID.NO.:67 SEQ.ID.NO.:69 SEQ.ID.NO.:71 SEQ.ID.NO.:73
GPR37 L352R C543Y	SEQ.ID.NO.:74 SEQ.ID.NO.:76	SEQ.ID.NO.:75 SEQ.ID.NO.:77
HF1948 I281F E135N	SEQ.ID.NO.:78 SEQ.ID.NO.:80	SEQ.ID.NO.:79 SEQ.ID.NO.:81
GPR66 T273K	SEQ.ID.NO.:82	SEQ.ID.NO.:83
GPR35 A216K	SEQ.ID.NO.:84	SEQ.ID.NO.:85
ETBR-LP2 N358K	SEQ.ID.NO.:86	SEQ.ID.NO.:87

Example 3**RECEPTOR EXPRESSION**

5

Although a variety of cells are available to the art-skilled for the expression of proteins, it is preferred that mammalian cells be utilized. The primary reason for this is predicated upon practicalities, *i.e.*, utilization of, *e.g.*, yeast cells for the expression of a GPCR, while possible, introduces into the protocol a non-mammalian cell which may not (indeed, in the case of yeast, does not) include the receptor-coupling, genetic-mechanism and secretory pathways that have evolved for mammalian systems — thus, results obtained in non-mammalian cells, while of potential use, are not as preferred as those obtained using mammalian cells. Of the mammalian cells, COS-7, 293 and 293T cells are particularly preferred, although the specific mammalian cell utilized can be predicated upon the particular needs of the artisan.

15

a. Transient Transfection of 293 Cells

On day one, 6×10^6 cells/10 cm dish of 293 cells well were plated out. On day two, two reaction tubes were prepared (the proportions to follow for each tube are per plate): tube A was prepared by mixing 4 μ g DNA (e.g., pCMV vector; pCMV vector with receptor cDNA, *etc.*) in 0.5 ml serum free DMEM (Gibco BRL); tube B was prepared by mixing 24 μ l lipofectamine (Gibco BRL) in 0.5ml serum free DMEM. Tubes A and B were admixed by inversion (several times), followed by incubation at room temperature for 30-45min. The admixture is referred to as the "transfection mixture". Plated 293 cells were washed with 1XPBS, followed by addition of 5 ml serum free DMEM. One ml of the transfection mixture were added to the cells, followed by incubation for 4hrs at 37°C/5% CO₂. The transfection mixture was removed by aspiration, followed by the addition of 10ml of DMEM/10% Fetal Bovine Serum. Cells were incubated at 37°C/5% CO₂. After 48hr incubation, cells were harvested and utilized for analysis.

b. Stable 293 Cell Lines

Approximately 12×10^6 293 cells will be plated on a 15cm tissue culture plate, and grown in DME High Glucose Medium containing 10% fetal bovine serum and one percent sodium pyruvate, L-glutamine, and antibiotics. Twenty-four hours following plating of 293 cells (to approximately ~80% confluency), the cells will be transfected using 12 μ g of DNA. The 12 μ g of DNA is combined with 60 μ l of lipofectamine and 2mL of DME High Glucose Medium without serum. The medium will be aspirated from the plates and the cells washed once with medium without serum. The DNA, lipofectamine, and medium mixture will be added to the plate along with 10mL of medium without serum. Following incubation at 37°C for four to five hours, the medium will be aspirated and 25ml of medium containing serum will be added. Twenty-four hours following transfection, the medium will be

aspirated again, and fresh medium with serum will be added. Forty-eight hours following transfection, the medium will be aspirated and medium with serum will be added containing geneticin (G418 drug) at a final concentration of 500µg/mL. The transfected cells will then undergo selection for positively transfected cells containing the G418 resistant gene. The medium will be replaced every four to five days as selection occurs. During selection, cells will be grown to create stable pools, or split for stable clonal selection.

C. RGT CELLS (USED FOR MGLUR7)

RGT cells were derived from an adenovirus transformed Syrian hamster cell line (AV12-664) into which a glutamate-aspartate transporter was stably transfected.

On day one, 5×10^6 / 10 cm dish of RGT cells were plated out. On day two, 91µl of serumfree media was added to a tube, followed by the addition of 9µl of Fugene 6 (Roche). To the same mix 3 ug of DNA was added (at 0.5 ug/ul). The mixture was gently mixed and incubated at room temperature for 15 min, then this mixture was added dropwise to the cells growing in DMEM/10% FBS and incubated for 48 hours at 37°C/5% CO₂. After 48hr incubation, cells were harvested and utilized for analysis.

Example 4

**ASSAYS FOR DETERMINATION OF CONSTITUTIVE ACTIVITY
OF NON-ENDOGENOUS GPCRS**

A variety of approaches are available for assessment of constitutive activity of the non-endogenous human GPCRs. The following are illustrative; those of ordinary skill in the art are credited with the ability to determine those techniques that are preferentially beneficial for the needs of the artisan.

1. Membrane Binding Assays: [³⁵S]GTPγS Assay

When a G protein-coupled receptor is in its active state, either as a result of ligand binding or constitutive activation, the receptor couples to a G protein and stimulates the

release of GDP and subsequent binding of GTP to the G protein. The alpha subunit of the G protein-receptor complex acts as a GTPase and slowly hydrolyzes the GTP to GDP, at which point the receptor normally is deactivated. Constitutively activated receptors continue to exchange GDP for GTP. The non-hydrolyzable GTP analog, [³⁵S]GTPγS, can be utilized to demonstrate enhanced binding of [³⁵S]GTPγS to membranes expressing constitutively activated receptors. Advantages of using [³⁵S]GTPγS binding to measure constitutive activation include but are not limited to the following: (a) it is generically applicable to all G protein-coupled receptors; (b) it is proximal at the membrane surface making it less likely to pick-up molecules which affect the intracellular cascade.

10 The assay takes advantage of the ability of G protein coupled receptors to stimulate [³⁵S]GTPγS binding to membranes expressing the relevant receptors. The assay can, therefore, be used in the direct identification method to screen candidate compounds to constitutively activated G protein-coupled receptors. The assay is generic and has application to drug discovery at all G protein-coupled receptors.

15 The [³⁵S]GTPγS assay is incubated in 20 mM HEPES and between 1 and about 20mM MgCl₂ (this amount can be adjusted for optimization of results, although 20mM is preferred) pH 7.4, binding buffer with between about 0.3 and about 1.2 nM [³⁵S]GTPγS (this amount can be adjusted for optimization of results, although 1.2 is preferred) and 12.5 to 75 μg membrane protein (*e.g.*, 293 cells expressing the G_s Fusion Protein; this amount can be adjusted for optimization) and 10 μM GDP (this amount can be changed for optimization) for 1 hour. Wheatgerm agglutinin beads (25 μl; Amersham) will then be added and the mixture incubated for another 30 minutes at room temperature. The tubes will be then centrifuged at 1500 x g for 5 minutes at room temperature and then counted in a scintillation counter.

20

2. Cell-based cAMP Detection Assay

A Flash Plate™ Adenylyl Cyclase kit (New England Nuclear, Cat. No. SMP004A) designed for cell-based assays can be modified for use with crude plasma membranes. The Flash Plate wells can contain a scintillant coating which also contains a specific antibody recognizing cAMP. The cAMP generated in the wells can be quantitated by a direct competition for binding of radioactive cAMP tracer to the cAMP antibody. The following serves as a brief protocol for the measurement of changes in cAMP levels in whole cells that express the receptors.

Transfected cells were harvested approximately twenty four hours after transient transfection. Media was carefully aspirated and discarded. Ten ml of PBS was gently added to each dish of cells followed by careful aspiration. One ml of Sigma cell dissociation buffer and 3ml of PBS was added to each plate. Cells were pipetted off the plate and the cell suspension collected into a 50ml conical centrifuge tube. Cells were centrifuged at room temperature at 1,100 rpm for 5 min. The cell pellet was carefully resuspended into an appropriate volume of PBS (about 3ml/plate). The cells were then counted using a hemocytometer and additional PBS was added to give the appropriate number of cells (to a final volume of about 50µl/well).

cAMP standards and Detection Buffer (comprising 1 µCi of tracer [¹²⁵I cAMP (50 µl) to 11 ml Detection Buffer) was prepared and maintained in accordance with the manufacturer's instructions. Assay Buffer was prepared fresh for screening and contained 50µl of Stimulation Buffer, 3µl of test compound (12µM final assay concentration) and 50µl cells, Assay Buffer was be stored on ice until utilized. The assay was initiated by addition of 50µl of cAMP standards to appropriate wells followed by addition of 50µl of PBSA to wells H-11 and H12. Fifty µl of Stimulation Buffer was added to all wells.

DMSO (or selected candidate compounds) was added to appropriate wells using a pin tool capable of dispensing 3 μ l of compound solution, with a final assay concentration of 12 μ M test compound and 100 μ l total assay volume. The cells were then added to the wells and incubated for 60 min at room temperature. One hundred μ l of Detection Mix containing
5 tracer cAMP was then added to the wells. Plates were incubated for an additional 2 hours followed by counting in a Wallac MicroBeta™ scintillation counter. Values of cAMP/well were then extrapolated from a standard cAMP curve which were contained within each assay plate.

10 3. Co-Transfection of Gi Coupled FPRL-2 with a Gs/Gi Fusion Protein Construct

The transfection mixture (from Example 3A) containing FPRL-2 and Gs/Gi Fusion Protein Construct was removed by aspiration, followed by the addition of 10ml of DMEM/10% Fetal Bovine Serum. Cells were then incubated at 37°C/5% CO₂. After
15 48hr incubation, cells were harvested and utilized for analysis. Cell-based cAMP detection assay was then performed according to the protocol in Example 4(2) above.

Because endogenous FPRL-2 is believed to predominantly couple with the Gi protein in its active state, a decrease in cAMP production signifies that the disclosed non-endogenous version of FPRL-2 is constitutively active. Thus, a candidate compound which
20 impacts the FPRL-2 receptor by increasing the cAMP signal is an inverse agonist, while a FPRL-2 agonist will decrease the cAMP signal. See, Figure 1.

Figure 1 evidence about a 4 fold increase in activity of FPRL-2 when compared to the Gs/Gi. When comparing the endogenous version of FPRL-2 with that of the non-endogenous version, the non-endogenous FPRL-2 ("FPRL-2(L240K)") evidence about a 3
25 fold increase in receptor activity when compared to the control, Gs/Gi. Therefore, this data

suggests that both the endogenous and non-endogenous versions of FPRL-2 are constitutively active.

Reference is made to Figure 9. In Figure 9, non-endogenous GPR37(L352R) produced about a 354% increase in cAMP when compared with the endogenous version of GPR37 ("GPR37 wt"), while GPR37(C543Y) produced about a 189% increase in cAMP when compared with GPR37 wt. This data suggests that both non-endogenous L352R and C543Y versions of GPR37 are constitutively activated.

4. Cell-Based cAMP for G_i Coupled Target GPCRs

TSHR is a G_s coupled GPCR that causes the accumulation of cAMP upon activation. TSHR will be constitutively activated by mutating amino acid residue 623 (*i.e.*, changing an alanine residue to an isoleucine residue). A G_i coupled receptor is expected to inhibit adenylyl cyclase, and, therefore, decrease the level of cAMP production, which can make assessment of cAMP levels challenging. An effective technique for measuring the decrease in production of cAMP as an indication of constitutive activation of a G_i coupled receptor can be accomplished by co-transfecting, most preferably, non-endogenous, constitutively activated TSHR (TSHR-A623I) (or an endogenous, constitutively active G_s coupled receptor) as a "signal enhancer" with a G_i linked target GPCR to establish a baseline level of cAMP. Upon creating a non-endogenous version of the G_i coupled receptor, this non-endogenous version of the target GPCR is then co-transfected with the signal enhancer, and it is this material that can be used for screening. This approach will be utilized to effectively generate a signal when a cAMP assay is used; this approach is preferably used in the direct identification of candidate compounds against G_i coupled receptors. It is noted that for a G_i coupled GPCR, when this approach is used, an inverse

agonist of the target GPCR will increase the cAMP signal and an agonist will decrease the cAMP signal.

Cells were transfected according to Example 3A above. The transfected cells were then transfected cells will be harvested approximately twenty four hours after transient
5 transfection. Cell-based cAMP detection assay was then performed according to the protocol in Example 4(2) above.

Preferably, and as noted previously, to ensure that a small molecule candidate compound is targeting the Gi coupled target receptor and not, for example, the TSHR(A623I), the directly identified candidate compound is preferably screened against
10 the signal enhancer in the absence of the target receptor.

Reference is made to Figure 3. Figure 3 is a comparative analysis of endogenous GPR45 ("GPR45 wt") versus a control ("CMV") in 293 cells. Endogenous target receptor GPR45 was co-transfected with a signal enhancer, TSHR(A623I). In the absence of TSH, the endogenous ligand for TSH receptor, co-transfection of TSHR(A623I) with endogenous
15 GPR45 evidence about a 96% decrease in production of cAMP when compared with the control (CMV). In the presence of TSH, endogenous GPR45 ("GPR45 wt") evidence about a 73% decrease in cAMP production when compared to the control ("CMV"). This data indicates that GPR45 is endogenously constitutively active and couples through the Gi protein.

20 Reference is made to Figure 4 and Table E. Table E is a summary of Figure 4, which is a comparative analysis of endogenous mGluR7 ("mGluR7 wt") with several non-endogenous versions of mGluR7 ("W590S," "R659H," "T771C" and "T790K") and the control ("pCMV") in 293 cells. Table E summarizes the cAMP production of the vector containing the signal enhancer receptor (*i.e.*, TSHR(A623I)) with the target receptor

(mGluR7) in the absence of its endogenous ligand (*i.e.*, TSH); the cAMP production of the co-transfection of the signal enhancer with the target receptor in the presence of TSH percent (%) decrease, in cAMP production, between the endogenous version of mGluR7 and the non-endogenous versions of mGluR7, co-transfected with TSHR(A623I) in the presence of TSH. This data evidences that the non-endogenous versions of mGluR7 ("W590S," "R659H," "T771C" and "I790K") reduce the production of cAMP when compared to the endogenous mGluR7, and thus has been constitutively activated by the methods disclosed above.

TABLE E

Versions of mGluR7	Co-Transfection of 1) Vector-TSHR(A623I) 2) mGluR7 versions 3) without 16mU/ml TSH (pmol cAMP)	Co-Transfection of 1) Vector-TSHR(A623I) 2) mGluR7 versions 3) 16mU/ml TSH (pmol cAMP)	Percent (%) Decrease between Endogenous and Non-endogenous Version of mGluR7 (with TSH)	mGluR7 Inverse Agonist	mGluR7 Agonist
pCMV (without TSHR)	4	--	--	Increase	Decrease
pCMV	23	288	--		
MgluR7 wt	21	402	0		
W590S	9	138	66		
R659H	7	156	61		
T771C	7	156	61		
I790K	9	151	62		

10

Versions of mGluR7 transfected in RGT cells support the data of above. Reference is made to Figure 5. In Figure 5, W590S evidenced about a 52% decrease in cAMP production; R659H evidenced about a 43% reduction; T771C evidenced about a 5% reduction; and I790K evidenced about a 28% reduction in the production of cAMP when compared to the endogenous version of mGluR7 receptor.

15

Because mGluR7 predominantly couples with Gi in its active state, a decrease in cAMP production signifies that the disclosed non-endogenous versions of mGluR7 are constitutively active. Thus, a candidate compound which impacts the mGluR7 receptor by increasing the cAMP signal is an inverse agonist, while a mGluR7 agonist will decrease the cAMP signal. Based upon the data generated for Figures 5 and 6, "W590S," "R659H," "T771C" and "I790K" are preferred non-endogenous versions of mGluR7, most preferably is "W590S" when used in a TSHR constitutively activated co-transfection approach using a cAMP assay in both 293 and RGT cells.

Reference is made to Figure 12. In Figure 12, non-endogenous versions of HF1948 ("I281F" and "E135N") evidenced a reduction in cAMP production, about an 18% and about a 39% reduction, respectively, when compared to the endogenous version of HF1948 ("wt"). This data suggests that both non-endogenous I281F and E135N versions of HF1948 are constitutively activated. This decrease in cAMP further suggests that these versions may be Gi-coupled. In addition to being Gi-coupled, Figure 11 suggests that non-endogenous I281F version of HF1948 may also couple to Gq G protein. (See, Example 4(5)(f) below).

Reference is made to Figure 16. Figure 16 evidences about a 36% decrease in cAMP production of cells co-transfected with TSHR-A623I ("TSHR-A623I") (in the presence of TSH) and non-endogenous, constitutively activated ETBR-LP2 ("N358K") (65.96 pmole cAMP/well) compared to TSHR-A623I with endogenous ETBR-LP2 ("WT") (102.59 pmol cAMP/well). About a 77% and about a 65% decrease in production of cAMP was evidenced when comparing TSHR-A623I co-transfected with ETBR-LP2("N358K") and TSHR-A623I co-transfected with ETBR-LP2("WT") against TSHR-A623I co-transfected with pCMV (290.75 pmol cAMP/well), respectively. Preferably, this approach

is used for screening an inverse agonist, which would increase the signal, whereas an agonist should decrease the signal. To confirm that a small molecule binds ETBR-LP2 and not to the TSHR-A623I construct, the small molecule is preferably screened against the construct in the absence of ETBR-LP2.

5 5. Reporter-Based Assays

a. CRE-LUC Reporter Assay (G_s -associated receptors)

293 and 293T cells were plated-out on 96 well plates at a density of 2×10^4 cells per well and were transfected using Lipofectamine Reagent (BRL) the following day according to manufacturer instructions. A DNA/lipid mixture was prepared for each 6-
10 well transfection as follows: 260ng of plasmid DNA in 100 μ l of DMEM are gently mixed with 2 μ l of lipid in 100 μ l of DMEM (the 260ng of plasmid DNA consisted of 200ng of a 8xCRE-Luc reporter plasmid, 50ng of pCMV comprising endogenous receptor or non-endogenous receptor or pCMV alone, and 10ng of a GPRS expression plasmid (GPRS in pcDNA3 (Invitrogen)). The 8XCRE-Luc reporter plasmid is prepared
15 as follows: vector SRIF- β -gal was obtained by cloning the rat somatostatin promoter (-71/+51) at BglV-HindIII site in the p β gal-Basic Vector (Clontech). Eight (8) copies of cAMP response element were obtained by PCR from an adenovirus template AdpCF126CCRE8 (*see, 7 Human Gene Therapy* 1883 (1996)) and cloned into the SRIF- β -gal vector at the Kpn-BglV site, resulting in the 8xCRE- β -gal reporter vector. The
20 8xCRE-Luc reporter plasmid was generated by replacing the beta-galactosidase gene in the 8xCRE- β -gal reporter vector with the luciferase gene obtained from the pGL3-basic vector (Promega) at the HindIII-BamHI site. Following 30 min. incubation at room temperature, the DNA/lipid mixture was diluted with 400 μ l of DMEM and 100 μ l of the diluted mixture was added to each well. One hundred μ l of DMEM with 10% FCS was

added to each well after a 4hr incubation in a cell culture incubator. The following day the transfected cells were changed with 200 μ l/well of DMEM with 10% FCS. Eight hours later, the wells were changed to 100 μ l /well of DMEM without phenol red, after one wash with PBS. Luciferase activity was measured the next day using the LucLite™ reporter gene assay kit (Packard) following manufacturer's instructions and read on a 1450 MicroBeta™ scintillation and luminescence counter (Wallac).

Reference is made to Figure 2. Figure 2 evidences about a 50% decrease in activity of STRL33 when compared to the control (CMV) at 12.5ng of STRL33 receptor. When comparing the endogenous version of STRL33 with that of the non-endogenous version, the non-endogenous STRL33 ("STRL33(L230K)") evidence about a 30% decrease in receptor activity when comparing at 12.5ng of protein, and about a 40% decrease in activity at 25 ng of protein. This data suggests that non-endogenous version of STRL33 receptor is constitutively active and may couple to the G protein, Gi.

b. AP1 reporter assay (G_q -associated receptors)

A method to detect G_q stimulation depends on the known property of G_q -dependent phospholipase C to cause the activation of genes containing AP1 elements in their promoter. A Pathdetect™ AP-1 cis-Reporting System (Stratagene, Catalogue # 219073) was utilized following the protocol set forth above with respect to the CREB reporter assay, except that the components of the calcium phosphate precipitate were 410 ng pAP1-Luc, 80 ng pCMV-receptor expression plasmid, and 20 ng CMV-SEAP.

Reference is made to Figure 17. Figure 17 represents a 61.1% increase in activity of the non-endogenous, constitutively active version of human ETBR-LP2 ("N358K") (2203 relative light units) compared with that of the endogenous ETBR-LP2 (862 relative

light units). This data suggests that non-endogenous version of ETBR-LP2 receptor is constitutively active and may couple to the G protein, Gi.

c. SRF-LUC Reporter Assay (G_q - associated receptors)

One method to detect G_q stimulation depends on the known property of G_q -
5 dependent phospholipase C to cause the activation of genes containing serum response factors in their promoter. A Pathdetect™ SRF-Luc-Reporting System (Stratagene) can be utilized to assay for G_q coupled activity in, *e.g.*, COS7 cells. Cells are transfected with the plasmid components of the system and the indicated expression plasmid encoding endogenous or non-endogenous GPCR using a Mammalian Transfection™ Kit
10 (Stratagene, Catalogue #200285) according to the manufacturer's instructions. Briefly, 410 ng SRF-Luc, 80 ng pCMV-receptor expression plasmid and 20 ng CMV-SEAP (secreted alkaline phosphatase expression plasmid; alkaline phosphatase activity is measured in the media of transfected cells to control for variations in transfection efficiency between samples) are combined in a calcium phosphate precipitate as per the
15 manufacturer's instructions. Half of the precipitate is equally distributed between 3 wells in a 96-well plate, kept on the cells in a serum free media for 24 hours. The last 5 hours the cells are incubated with 1 μ M Angiotensin, where indicated. Cells are then lysed and assayed for luciferase activity using a Lucite™ Kit (Packard, Cat. # 6016911) and "Trilux 1450 Microbeta" liquid scintillation and luminescence counter (Wallac) as
20 per the manufacturer's instructions. The data can be analyzed using GraphPad Prism™ 2.0a (GraphPad Software Inc.).

d. SRE Reporter Assay

A SRE-Luc Reporter (a component of Mercury Luciferase System 3, Clontech
25 Catalogue # K2053-1) was utilized in 293 cells. Cells were transfected with the plasmid

components of this system and the indicated expression plasmid encoding endogenous or non-endogenous receptor using Lipofectamine Reagent (Gibco/BRL, Catalogue #18324-012) according to the manufacturer's instructions. Briefly, 420ng SRE-Luc, 50ng CMV (comprising the GPR37 receptor) and 30 ng CMV-SEAP (secreted alkaline phosphatase expression plasmid; alkaline phosphatase activity is measured in the media of transfected cells to control for variations in transfection efficiency between samples) were combined in a cationic lipid-DNA precipitate as per the manufacturer's instructions. The final volume was 25µl brought up with Optimem (Vendor). This is referred to as the "template mix." The template mix was combined with the lipfectamine in a polystyrene tube and was incubated for 30 minutes. During the incubation, the cells were washed with 100µl Optimem. After incubation, 200µl of Optimem was added to mix and 40µl-50µl/well. The cells were left to mix overnight. The media was replaced with fresh medium the following morning to DMEM/Phenol red free/1% FBNS at 130µl/well. The cells were then assayed for luciferase activity using a Lucite™ Kit (Packard, Cat. # 6016911) and "Trilux 1450 Microbeta" liquid scintillation and luminescence counter (Wallac) as per the manufacturer's instructions. The data were analyzed using GraphPad Prism™ 2.0a (GraphPad Software Inc.).

Reference is made to Figure 7. In Figure 7, when comparing the non-endogenous version of GPR37 ("C543Y") with the endogenous version ("wt"), the C543Y mutation evidences about a 316% increase in cAMP production over the wt version, while the non-endogenous version "L352R" evidence about a 178% increase in production of cAMP over the wt version. This data suggests that both non-endogenous versions of GPR37, C543Y and L352R, are constitutively activated.

e. **E2F-Luc Reporter Assay**

A pE2F-Luc Reporter (a component of Mercury Luciferase System 3, Clontech Catalogue # K2053-1) was utilized in 293A cells. Cells were transfected with the plasmid components of this system and the indicated expression plasmid encoding endogenous or non-endogenous receptor using Lipofectamine Reagent (Gibco/BRL, Catalogue #18324-012) according to the manufacturer's instructions. Briefly, 400 ng pE2F-Luc, 80 ng CMV (comprising the GPR35 receptor) and 20 ng CMV-SEAP (secreted alkaline phosphatase expression plasmid; alkaline phosphatase activity is measured in the media of transfected cells to control for variations in transfection efficiency between samples) were combined in a cationic lipid-DNA precipitate as per the manufacturer's instructions. Half of the precipitate was equally distributed over 3 wells in a 96-well plate, kept on the cells overnight, and replaced with fresh medium the following day. Forty-eight (48) hr after the start of the transfection, cells were treated and assayed for luciferase activity using a Lucite™ Kit (Packard, Cat. # 6016911) and "Trilux 1450 Microbeta" liquid scintillation and luminescence counter (Wallac) as per the manufacturer's instructions. The data were analyzed using GraphPad Prism™ 2.0a (GraphPad Software Inc.).

Reference is made to Figure 14. Figure 14 represents about a 100% increase in activity of the non-endogenous, constitutively active version of human GPR35 (A216K) (607.13 relative light units) compared with that of the endogenous GPR35 (24.97 relative light units). This data suggests that GPR35(A216K) interacts with the transcription factor E2F to drive the expression of the luciferase protein. Such interaction with E2F, along with evidence that GPR35 is expressed in colorectal cancer cells, further suggests that GPR35 may play a role in cancer cell proliferation. Thus, based upon these data, a preferred candidate compound which impacts the GPR35 receptor would be an inverse agonist. This

data suggest that an inverse agonist of GPR35 would be useful in the treatment of cancerous conditions, colorectal cancer in particular.

f. Intracellular IP₃ Accumulation Assay (G_q-associated receptors)

5 On day 1, cells comprising the receptors (endogenous and/or non-endogenous) are plated onto 24 well plates, usually 1×10^5 cells/well (although this number can be optimized. On day 2 cells were transfected by firstly mixing 0.25ug DNA in 50 μ l serum free DMEM/well and 2 μ l lipofectamine in 50 μ l serum free DMEM/well. The solutions were gently mixed and incubated for 15-30 min at room temperature. Cells were then washed
10 with 0.5 ml PBS and 400 μ l of serum free media and then mixed with the transfection media and added to the cells. The cells were incubated for 3-4 hrs at 37°C/5%CO₂ and then the transfection media was removed and replaced with 1ml/well of regular growth media. On day 3 the cells are labeled with ³H-myo-inositol. Briefly, the media was removed and the cells are washed with 0.5 ml PBS. Then 0.5 ml inositol-free/serum free media (GIBCO
15 BRL) were added/well with 0.25 μ Ci of ³H-myo-inositol/ well and the cells incubated for 16-18 hrs overnight at 37°C/5%CO₂. On Day 4 the cells are washed with 0.5 ml PBS and 0.45 ml of assay medium was added containing inositol-free/serum free media 10 μ M pargyline 10 mM lithium chloride or 0.4 ml of assay medium. The cells were then incubated for 30 min at 37°C. The cells are then washed with 0.5 ml PBS and 200 μ l of
20 fresh/ice cold stop solution (1M KOH; 18 mM Na-borate; 3.8 mM EDTA) is added to each well. The solution was kept on ice for 5-10 min (or until cells are lysed) and then neutralized by 200 μ l of fresh/ice cold neutralization solution (7.5 % HCL). The lysate was then transferred into 1.5 ml Eppendorf tubes and 1 ml of chloroform/methanol (1:2) was added/tube. The solution was vortexed for 15 sec and the upper phase was applied to a
25 Biorad AG1-X8™ anion exchange resin (100-200 mesh). First, the resin was washed with

water at 1:1.25 W/V and 0.9 ml of upper phase was loaded onto the column. The column was then washed with 10 ml of 5 mM myo-inositol and 10 ml of 5 mM Na-borate/60mM Na-formate. The inositol tris phosphates were eluted into scintillation vials containing 10 ml of scintillation cocktail with 2 ml of 0.1 M formic acid/ 1 M ammonium formate. The
5 columns were regenerated by washing with 10 ml of 0.1 M formic acid/3M ammonium formate and rinsed twice with dd H₂O and stored at 4°C in water.

Reference is made to Figure 6. In Figure 6, 293 cells were transfected with Gq protein containing a six amino acid deletion, "Gq(del)"; Gq protein fused to a Gi protein, "Gq(del)/Gi", and non-endogenous mGluR7, T771C together with Gq(del),
10 "T771C+Gq(del)" and T771C with Gq(del)/Gi, "T771C+Gq(del)/Gi". Inositol triphosphate was measured in the presence and absence of glutamate. Co-transfection of non-endogenous version of mGluR7 with Gq(del)/Gi evidence about a 1850 fold increase when compared to the Gq(del)/Gi in the presence of glutamate; and about a 860 fold increase compared with T771C+Gq(del)/Gi in the presence of glutamate. These data evidences that
15 mGluR7, a Gi coupled receptor, can be activated via the Gq protein. Therefore, the Gq(del)/Gi Fusion Construct can be co-transfected with a GPCR and used to as a tool to screen for candidate compounds.

Reference is made to Figure 11. In Figure 11, when comparing the non-endogenous version of HF1948 ("I281F") with the endogenous version ("wt"), the I281F mutation
20 evidences about a 361% increase in IP3 accumulation over the wt version. This data suggests that the non-endogenous I281F version of HF1948 is constitutively activated and is Gq-coupled.

Example 5

FUSION PROTEIN PREPARATION

25 a. GPCR: G_i Fusion Construct

The design of the constitutively activated GPCR-G protein fusion construct can be accomplished as follows: both the 5' and 3' ends of the rat G protein $G_s\alpha$ (long form; Itoh, H. et al., 83 *PNAS* 3776 (1986)) is engineered to include a HindIII (5'-AAGCTT-3') sequence thereon. Following confirmation of the correct sequence (including the flanking HindIII sequences), the entire sequence is shuttled into pcDNA3.1(-) (Invitrogen, cat. no. V795-20) by subcloning using the HindIII restriction site of that vector. The correct orientation for the $G_s\alpha$ sequence will be determined after subcloning into pcDNA3.1(-). The modified pcDNA3.1(-) containing the rat $G_s\alpha$ gene at HindIII sequence is then verified; this vector will then be available as a "universal" $G_s\alpha$ protein vector. The pcDNA3.1(-) vector contains a variety of well-known restriction sites upstream of the HindIII site, thus beneficially providing the ability to insert, upstream of the G_s protein, the coding sequence of an endogenous, constitutively active GPCR. This same approach can be utilized to create other "universal" G protein vectors, and, of course, other commercially available or proprietary vectors known to the artisan can be utilized. In some embodiments, the important criteria is that the sequence for the GPCR be upstream and in-frame with that of the G protein.

Spacers in the restriction sites between the G protein and the GPCR are optional. The sense and anti-sense primers included the restriction sites for XbaI and EcoRV, respectively, such that spacers (attributed to the restriction sites) exist between the G protein and the GPCR.

PCR will then be utilized to secure the respective receptor sequences for fusion within the $G_s\alpha$ universal vector disclosed above, using the following protocol for each: 100ng cDNA for GPCR will be added to separate tubes containing 2 μ l of each primer

(sense and anti-sense), 3µl of 10mM dNTPs, 10µl of 10XTaqPlus™ Precision buffer, 1µl of TaqPlus™ Precision polymerase (Stratagene: #600211), and 80µl of water. Reaction temperatures and cycle times for the GPCR will be as follows with cycle steps 2 through 4 were repeated 35 times: 94°C for 1 min; 94°C for 30 seconds; 62°C for 20 sec; 72°C 1 min
 5 40sec; and 72°C 5 min. PCR products will be run on a 1% agarose gel and then purified. The purified products will be digested with XbaI and EcoRV and the desired inserts purified and ligated into the G_s universal vector at the respective restriction sites. The positive clones will be isolated following transformation and determined by restriction enzyme digestion; expression using 293 cells will be accomplished following the protocol
 10 set forth *infra*. Each positive clone for GPCR- G_s Fusion Protein will be sequenced to verify correctness.

g. G_q(6 amino acid deletion)/G_i Fusion Construct

The design of a G_q(del)/G_i fusion construct was accomplished as follows: the N-terminal six (6) amino acids (amino acids 2 through 7), having the sequence of TLESIM
 15 (SEQ.ID.NO.:88) Gαq-subunit was deleted and the C-terminal five (5) amino acids, having the sequence EYNLV (SEQ.ID.NO.:89) was replaced with the corresponding amino acids of the Gαi Protein, having the sequence DCGLF (SEQ.ID.NO.:90). This fusion construct was obtained by PCR using the following primers:

5'-gatcAAGCTTCCATGGCGTGCTGCCTGAGCGAGG-3' (SEQ.ID.NO.:91) and
 20 5'-gatcGGATCCTTAGAACAGGCCGCAGTCCTTCAGGTTTCAGCTGCAGGATGGTG-3'
 (SEQ.ID.NO.:92) and Plasmid 63313 which contains the mouse Gαq-wild type version with a hemagglutinin tag as template. Nucleotides in lower caps are included as spacers.

TaqPlus® Precision DNA polymerase (Stratagene) was utilized for the amplification by the following cycles, with steps 2 through 4 repeated 35 times: 95°C for

2 min; 95°C for 20 sec; 56°C for 20 sec; 72°C for 2 min; and 72°C for 7 min. The PCR product will be cloned into a pCRII-TOPO vector (Invitrogen) and sequenced using the ABI Big Dye Terminator kit (P.E. Biosystems). Inserts from a TOPO clone containing the sequence of the fusion construct will be shuttled into the expression vector pcDNA3.1(+) at the HindIII/BamHI site by a 2 step cloning process.

c. Gs/Gi Fusion Protein Construct

The design of a Gs/Gi Fusion Protein Construct was accomplished as follows: the C-terminal five (5) amino acids of G α s-subunit was deleted, having the sequence 5'-QYELL-3' (SEQ.ID.NO.:93) and replaced with the corresponding amino acids of the G α i protein, having the sequence 5'-DCGLF-3' (SEQ.ID.NO.:94). This protein fusion construct was obtained by PCR using a 5' and 3' oligonucleotides.

TaqPlus Precision DNA polymerase (Stratagene) was utilized for the amplification by the following cycles, with steps 2 through 4 repeated 25 times: 98°C for 2 min; 98°C for 30 sec; 60°C for 30 sec; 72°C for 2 min; and 72°C for 5 min. The PCR product was cloned into a pCRII-TOPO vector (Invitrogen) and sequenced using the ABI Big Dye Terminator kit (P.E. Biosystems). Inserts from a TOPO clone containing the sequence of the protein fusion construct was shuttled into the expression vector pcDNA3.1(+) at the restriction site. The nucleic acid sequence for Gs/Gi Protein Fusion Construct was then determined. See SEQ.ID.NO.:95 for the nucleic acid sequence and SEQ.ID.NO.:96 for the amino acid sequence.

Example 6

SCHWANN CELL PREPARATION

2L of neonate rat pups (Sprague Dawley) (at Post-pardum day 2-Post-pardum day 3 stage) were placed on ice to euthanize. Pups were then removed and decapitated to drain

the blood. The neonates were placed, belly-down, on a dissection board and rinsed with 70% ethanol to sterilize. Using a scalpel, the skin was removed in the thigh area until the sciatic nerve was exposed (or until a thin white "string" extended from the spinal cord to the knee was visible). The nerves were placed in DMEM medium and then aspirated, followed
5 by bringing the volume to 2.4 ml with DMEM media and adding 300uL 10X Collagenase (0.3%, Sigma Cat. #C-9891) and 300uL 10X Trypsin (0.25%, GIBCO Cat. #25095-019) for dissociation. Nerves were then incubated at 37°C for 15 min, centrifuged for 5 min at 1,000 rpm followed by removing the media (repeated twice). 1 mL DMEM-HEPES and 1mL DMEM/10% FBS were added and then transferred to a 50mL conical tube. The contents of
10 the tube were sheared with the following gauge needles (VWR): once with 18G, twice with 21G and twice with 23G. The contents were placed on a Falcon cell strainer and spun at a very low speed (about 1200 rpm). The total volume was brought to 10mL with DMEM/10% FBS and plated on a Poly-L-lysine treated 10cm plate (Sigma, Cat. #P-1274). Plates were then incubated overnight in 37°C humid incubator at 7% CO₂. Fresh media
15 added with 100X ARA C (10mM, Sigma, Cat. #C-1768) and cultured for an additional 48 hours. The cells were then washed with PBS (three times) to remove the ARA C and the following were added: DMEM/10% FBS, different concentrations of Forskolin in 100% ethanol (2uM, 5uM, 10uM, 20uM and 50uM) (Calbiochem, Cat#344270), 80ug of Pituitary Extract (Sigma, #P-1167) in PBS and 0.1%BSA, followed by growing the cells for 30 hours
20 at 37°C humidifier at 7% CO₂. The cells were then collected and the RNA was isolated and analyzed.

Antibody selection was accomplished according to the following: the Poly-L-Lysine treated plates were first washed with 1X PBS (three times), trypsinized with 1mL of 0.5% trypsin-EDTA, for about 1 min and then neutralized with 9mL of DMEM-HEPES buffer

and 10% FBS. Cells were centrifuged at 1200rpm for 5 min, resuspended in 3mL of DMEM-HEPES to wash out the trypsin and spun for 5 min at 1200rpm. Cells were then resuspended in 600uL of DMEM-HEPES, leaving some media after the spin in order to have single cells. Thy1.1 antibody (Monoclonal Antibody, Sigma, Cat. #P-1274) was
5 added at a 1:1000 dilution.

The cells were then incubated for 20 min at 37°C, slightly agitating the tube every two minutes. 20uL of Guinea Pig complement (GIBCO, Cat. #19195-015) was thawed before using it, followed by adding the complement to the cells with the antibody to a final volume of 1mL. The cells were incubated for about 20 min-30 min at 37°C water bath and
10 10mL of DMEM-HEPES was added and spun down for 5 min at 1200rpm. Cells were resuspended in 5mLs of DMEM/10% FBS and added to poly-L-lysine treated plates that contains pituitary extract and forskolin. The cells were left to recover for 24-48 hours and the immune selection procedure was repeated twice.

EXAMPLE 7

15 PREPARATION OF CRUSHED RAT SCIATIC NERVE

The sciatic nerves of anesthetized (iso-florene), adult (10-13 week old) Sprague-Dawley rats were exposed at the sciatic notch. Nerve crush was produced by tightly compressing the sciatic nerve at the sciatic notch with flattened forceps twice, each time for 10 sec; this technique causes the axons to degenerate, but allows axonal regeneration.
20 At varying times after nerve injury, the animals were euthanized by CO₂ inhalation, the distal nerve stumps were removed, and the most proximal 2-3 mm was trimmed off. For crushed nerves, the entire distal nerve was harvested. The nerves were immediately frozen in liquid nitrogen and stored at -80°C. Unlesioned sciatic nerves were obtained from animals of varying ages, from P0 (post crush) to P13.

Example 8**TISSUE DISTRIBUTION OF THE DISCLOSED HUMAN GPCRS:****1. RT-PCR**

5

RT-PCR can be applied to confirm the expression and to determine the tissue distribution of several novel human GPCRs. Oligonucleotides utilized will be GPCR-specific and the human multiple tissue cDNA panels (MTC, Clontech) as templates. Taq DNA polymerase (Stratagene) will be utilized for the amplification in a 40 μ l reaction according to the manufacturer's instructions. Twenty μ l of the reaction will be loaded on a 1.5% agarose gel to analyze the RT-PCR products.

2. Dot-Blot

15

Using a commercially available human-tissue dot-blot format, endogenous GPCR was used to probe for a determination of the areas where such receptor is localized. The PCR fragments of Example 1 were used as the probe: radiolabeled probe was generated using this fragment and a Prime-It II™ Random Primer Labeling Kit (Stratagene, #300385), according to manufacturer's instructions. A human RNA Master Blot™ (Clontech, #7770-1) was hybridized with GPCR radiolabeled probe and washed under stringent conditions according manufacturer's instructions. The blot was exposed to Kodak BioMax Autoradiography film overnight at -80°C. Table F, below, lists the receptors and the tissues wherein expression was found. Exemplary diseases/disorders linked to the receptors are discussed in Example 6, *infra*.

25

TABLE F

Receptor Identifier	Tissue Expression
STRL33	Placenta, spleen and lung
GPR45	Central nervous system, brain
GPR37	central nervous system, specifically in the brain tissues, pituitary gland and placenta

GPR66	pancreas, bone, testis, mammary glands, small intestine, and spleen
GPR26	Brain
ETBR-LP2	Brain, pituitary gland and placenta

3. Northern Blot

a. GPR37

5 RNA from Example 6 was harvested utilizing RNazol B reagent (TelTest Inc., Cat. #CS-104), according to manufacturer's instructions. After electrophoresis in an 1% agarose/formaldehyde gel, the RNA was transferred to a nylon membrane (Sachleicher Schull) by capillary action using 10X SSC. A ³²P-labelled GPR37 DNA probe was synthesized using a DNA fragment corresponding precisely to the 3' end of GPR37 and a

10 High Prime labeling kit (Roche Molecular Biochemical) according to the manufacturer's instructions. Hybridization was performed using ExpressHyb Solution (Clontech, Cat. #8015-2) supplemented with 100 µg/ml salmon sperm DNA as follows. The membrane containing the separated RNA samples was first incubated with ExpressHyb solution at 65°C overnight. The ³²P-labelled GPR37 DNA probe was denatured by boiling for 2

15 minutes, placed on ice for 5 minutes and then transferred into the ExpressHyb solution bathing the membrane. After an overnight incubation at 65°C, the membrane was removed from the hybridization solution and washed four times for 15 minutes each in 2XSSC/1% SDS at 65°C, followed by two washes for 15 minutes each in 0.2XSSC/0.1% SDS at 55°C. Excess moisture was removed from the blot by gentle shaking, after which the blot was

20 wrapped in plastic wrap and exposed to film overnight at -80°C.

Reference is made to Figure 9. Figure 9 evidences that GPR37 is expressed in Schwann cells, such that myelination can be maintained at 20uM Forskolin.

Figure 10 evidences that GPR37 is up-regulated in crushed rat sciatic nerves, specifically seven (7) days after crushing the nerves. Such data is consistent with the data presented in Figure 9, *i.e.*, GPR37 may play a role in the regeneration of nerves by stimulating the process of myelination in Schwann cells.

5 GPR37 is expressed in the human central nervous system, specifically in the brain tissues. It has been further determined that GPR37 is expressed in Schwann cells. When axons (or nerves) are injured, Schwann cells act to regenerate the nerves by forming myelin sheaths around the axons, which provides "insulation" in the form of myelin sheaths. This process, known as myelination, is important in that action potentials travel at a faster rate,
10 thereby conserving metabolic energy. Schwann cells and their precursors play an important role in influencing the survival and differentiation of other cells that make up a peripheral nerve. In addition, GPR37 has been determined to be expressed in crushed rat sciatic nerves. Such data supports the evidence that GPR37 may play a role in regenerating nerve cells. Based on the known functions of the specific tissues to which the receptor is
15 localized, the putative functional role of the receptor can be deduced. Thus, in the case of hyper-myelination (*e.g.*, tumorigenesis), an inverse agonist against GPR37 is preferred, while an agonist is preferred where hypo-myelination occurs (*e.g.*, a degenerative disease such as diabetes).

b. GPR66

20 Total RNA from several pancreatic cell lines (*e.g.*, HIT, ARIP, Tu6, RIN α TC, STC, NIT, and EcR-CHO, all of which are commercially available) were isolated using TRIzol reagent (Gibco/BRL, Cat #15596-018) according to manufacturer's instructions. After electrophoresis in a 1% agarose/formaldehyde gel, the RNA was transferred to a nylon membrane using standard protocols. A 32 P-labelled GPR66 probe was synthesized

using a DNA fragment corresponding precisely to the entire coding sequence and a Prime It II Random Primer Labeling Kit (Stratagene, Cat. #300385) according to manufacturer's instructions. Hybridization was performed using ExpressHyb Solution (Clontech, Cat.#8015-2) supplemented with 100ug/ml salmon sperm DNA as follows. The membrane
5 containing the separated RNA samples were first incubated with ExpressHyb solution at 65°C for 1 hour. The ³²P-labeled GPR66 DNA probe was denatured by boiling for 2 min, placed on ice for 5 min and then transferred into the ExpressHyb solution bathing the membrane. After an overnight incubation at 65°C, the membrane was removed from the hybridization and washed four times for 15 min each in 2XSSC/1% SDS at 65°C, followed
10 by two washes for 15 min each in 0.1XSSC/0.5% SDS at 55°C. Excess moisture was removed from the blot by gentle shaking, after which the blot was wrapped in plastic and exposed to film overnight at -80°C.

Reference is made to Figure 13. Results of RNA blots (*see*, Figure 13) together with the dot-blot data, evidencing the expression of GPR66 in the pancreas, suggest that
15 GPR66 is abundantly expressed in all islet cell lines and in ARIP cells, a pancreatic ductal cell lines. While not wishing to be bound by any theory, the expression of GPR66 in the pancreatic cell lines suggest that GPR66 may play a role in islet neogenesis.

c. GPR35

Total RNA from several cancer cell lines (*e.g.*, RIN-5AH, HEP-G2, A549,
20 HELA, MOLT-4, HL-60 and SW480 cells, all of which are commercially available) were isolated using TRIzol reagent (Gibco/BRL, Cat #15596-018) according to manufacturer's instructions. After electrophoresis in a 1% agarose/formaldehyde gel, the RNA was transferred to a nylon membrane using standard protocols. A ³²P-labelled GPR35 probe was synthesized using a DNA fragment corresponding precisely to the

entire coding sequence and a Prime It II Random Primer Labeling Kit (Stratagene, Cat. #300385) according to manufacturer's instructions. Hybridization was performed using ExpressHyb Solution (Clontech, Cat.#8015-2) supplemented with 100ug/ml salmon sperm DNA as follows. The membrane containing the separated RNA samples were
5 first incubated with ExpressHyb solution at 65°C for 1 hour. The ³²P-labeled GPR35 DNA probe was denatured by boiling for 2 min, placed on ice for 5 min and then transferred into the ExpressHyb solution bathing the membrane. After an overnight incubation at 65°C, the membrane was removed from the hybridization and washed four times for 15 min each in 2XSSC/1% SDS at 65°C, followed by two washes for 15 min
10 each in 0.1XSSC/0.5% SDS at 55°C. Excess moisture was removed from the blot by gentle shaking, after which the blot was wrapped in plastic and exposed to film overnight at
-80°C.

Reference is made to Figure 15. Results of RNA blots (*see*, Figure 15) evidences
15 that GPR35 is abundantly expressed in colorectal cancer cell line SW480. Such data suggests that GPR35 may play a role in colorectal carcinogenesis. Identification of candidate compounds, by the method discussed below, is most preferably an inverse agonist. An inverse agonist for GPR35 is intended to reduce DNA replication in an effort to inhibit cell proliferation of cancerous cells. GPR35 is expressed in large and small
20 intestine. Numerous cancer cell lines were examined where GPR35 was determined to be expressed in the colorectal cancer cell line (e.g., HELA, MOLT-4 and SW480). This data suggests that GPR35 may play a role in colorectal carcinogenesis. Colorectal cancer is a malignancy that arises from either the colon or the rectum. Cancers of the large intestine are the second most common form of cancer found in both males and females.

d. ETBR-LP2

RNA from Example 6 was harvested utilizing RNazol B reagent (TelTest Inc., Cat. #CS-104), according to manufacturer's instructions. After electrophoresis in an 1% agarose/formaldehyde gel, the RNA was transferred to a nylon membrane (Sachleicher
5 Schull) by capillary action using 10X SSC. A ^{32}P -labelled ETBR-LP2 DNA probe was synthesized using a DNA fragment corresponding precisely to the 3' end of ETBR-LP2 and a High Prime labeling kit (Roche Molecular Biochemical) according to the manufacturer's instructions. Hybridization was performed using ExpressHyb Solution (Clontech, Cat. #8015-2) supplemented with 100 $\mu\text{g}/\text{ml}$ salmon sperm DNA as follows. The membrane
10 containing the separated RNA samples was first incubated with ExpressHyb solution at 65°C overnight. The ^{32}P -labelled ETBR-LP2 DNA probe was denatured by boiling for 2 minutes, placed on ice for 5 minutes and then transferred into the ExpressHyb solution bathing the membrane. After an overnight incubation at 65°C, the membrane was removed from the hybridization solution and washed four times for 15 minutes each in 2XSSC/1%
15 SDS at 65°C, followed by two washes for 15 minutes each in 0.2XSSC/0.1% SDS at 55°C. Excess moisture was removed from the blot by gentle shaking, after which the blot was wrapped in plastic wrap and exposed to film overnight at -80°C.

Reference is made to Figure 18. Figure 18 evidences that ETBR-LP2 is expressed in Schwann cells, such that myelination can be maintained at 20uM Forskolin.

20 Reference is made to Figure 19. Figure 19 evidences that ETBR-LP2 is up-regulated in crushed rat sciatic nerves, specifically seven (7) days after crushing the nerves. Such data is consistent with the data presented in Figure 18, *i.e.*, ETBR-LP2 may play a role in the regeneration of nerves by stimulating the process of myelination in Schwann cells.

Based upon these data, ETBR-LP2 is expressed in Schwann cells. When axons (or nerves) are injured, Schwann cells act to regenerate the nerves by forming myelin sheaths around the axons, which provides "insulation" in the form of myelin sheaths. This process, known as myelination, is important in that action potentials travel at a faster rate, thereby conserving metabolic energy. Schwann cells and their precursors play an important role in influencing the survival and differentiation of other cells that make up a peripheral nerve. In addition, ETBR-LP2 has been determined to be expressed in crushed rat sciatic nerves. Such data supports the evidence that ETBR-LP2 may play a role in regenerating nerve cells. Based on the known functions of the specific tissues to which the receptor is localized, the putative functional role of the receptor can be deduced. Thus, in the case of hypermyelination (*e.g.*, tumorigenesis), an inverse agonist against ETBR-LP2 is preferred, while an agonist is preferred where hypo-myelination occurs (*e.g.*, a degenerative disease such as diabetes).

Diseases and disorders related to receptors located in these tissues or regions include, but are not limited to, cardiac disorders and diseases (*e.g.* thrombosis, myocardial infarction; atherosclerosis; cardiomyopathies); kidney disease/disorders (*e.g.*, renal failure; renal tubular acidosis; renal glycosuria; nephrogenic diabetes insipidus; cystinuria; polycystic kidney disease); eosinophilia; leukocytosis; leukopenia; ovarian cancer; sexual dysfunction; polycystic ovarian syndrome; pancreatitis and pancreatic cancer; irritable bowel syndrome; colon cancer; Crohn's disease; ulcerative colitis; diverticulitis; Chronic Obstructive Pulmonary Disease (COPD); Cystic Fibrosis; pneumonia; pulmonary hypertension; tuberculosis and lung cancer; Parkinson's disease; movement disorders and ataxias; learning and memory disorders; eating disorders (*e.g.*, anorexia; bulimia, etc.); obesity; cancers; thymoma; myasthenia gravis; circulatory disorders; prostate cancer;

prostatitis; kidney disease/disorders(e.g., renal failure; renal tubular acidosis; renal glycosuria; nephrogenic diabetes insipidus; cystinuria; polycystic kidney disease); sensorimotor processing and arousal disorders; obsessive-compulsive disorders; testicular cancer; priapism; prostatitis; hernia; endocrine disorders; sexual dysfunction; allergies; 5 depression; psychotic disorders; migraine; reflux; schizophrenia; ulcers; bronchospasm; epilepsy; prostatic hypertrophy; anxiety; rhinitis; angina; and glaucoma. Accordingly, the methods of the present invention may also be useful in the diagnosis and/or treatment of these and other diseases and disorders.

10 **Example 7**

Protocol: Direct Identification of Inverse Agonists and Agonists

A. [³⁵S]GTPγS Assay

Although endogenous, constitutively active GPCRs have been used for the direct identification of candidate compounds as, e.g., inverse agonists, for reasons that are not 15 altogether understood, intra-assay variation can become exacerbated. In some embodiments a GPCR Fusion Protein, as disclosed above, is also utilized with a non-endogenous, constitutively activated GPCR. When such a protein is used, intra-assay variation appears to be substantially stabilized, whereby an effective signal-to-noise ratio is obtained. This has the beneficial result of allowing for a more robust identification of 20 candidate compounds. Thus, in some embodiments it is preferred that for direct identification, a GPCR Fusion Protein be used and that when utilized, the following assay protocols be utilized.

1. Membrane Preparation

Membranes comprising the constitutively active orphan GPCR Fusion Protein of interest and for use in the direct identification of candidate compounds as inverse agonists or agonists are preferably prepared as follows:

a. Materials

5 “Membrane Scrape Buffer” is comprised of 20mM HEPES and 10mM EDTA, pH 7.4; “Membrane Wash Buffer” is comprised of 20 mM HEPES and 0.1 mM EDTA, pH 7.4; “Binding Buffer” is comprised of 20mM HEPES, 100 mM NaCl, and 10 mM MgCl₂, pH 7.4

b. Procedure

10 All materials will be kept on ice throughout the procedure. Firstly, the media will be aspirated from a confluent monolayer of cells, followed by rinse with 10ml cold PBS, followed by aspiration. Thereafter, 5ml of Membrane Scrape Buffer will be added to scrape cells; this will be followed by transfer of cellular extract into 50ml centrifuge tubes (centrifuged at 20,000 rpm for 17 minutes at 4°C). Thereafter, the supernatant will be
15 aspirated and the pellet will be resuspended in 30ml Membrane Wash Buffer followed by centrifugation at 20,000 rpm for 17 minutes at 4°C. The supernatant will then be aspirated and the pellet resuspended in Binding Buffer. The resuspended pellet will then be homogenized using a Brinkman Polytron™ homogenizer (15-20 second bursts until the material is in suspension). This is referred to herein as “Membrane Protein”.

20 **2. Bradford Protein Assay**

Following the homogenization, protein concentration of the membranes will be determined, for example, using the Bradford Protein Assay (protein can be diluted to about 1.5mg/ml, aliquoted and frozen (-80°C) for later use; when frozen, protocol for use will be as follows: on the day of the assay, frozen Membrane Protein is thawed at room

temperature, followed by vortex and then homogenized with a Polytron at about 12 x 1,000 rpm for about 5-10 seconds; it was noted that for multiple preparations, the homogenizer is thoroughly cleaned between homogenization of different preparations).

a. Materials

5 Binding Buffer (as discussed above); Bradford Dye Reagent; Bradford Protein Standard will be utilized, following manufacturer instructions (Biorad, cat. no. 500-0006).

b. Procedure

Duplicate tubes will be prepared, one including the membrane, and one as a
10 control "blank". Each contains 800 μ l Binding Buffer. Thereafter, 10 μ l of Bradford Protein Standard (1mg/ml) will be added to each tube, and 10 μ l of membrane Protein will then be added to just one tube (not the blank). Thereafter, 200 μ l of Bradford Dye Reagent will be added to each tube, followed by vortexing. After five minutes, the tubes will be re-vortexed and the material therein will be transferred to cuvettes. The cuvettes
15 will then be read using a CECIL 3041 spectrophotometer, at wavelength 595.

3. Direct Identification Assay

a. Materials

GDP Buffer consisted of 37.5 ml Binding Buffer and 2mg GDP (Sigma, cat. no. G-7127), followed by a series of dilutions in Binding Buffer to obtain 0.2 μ M GDP (final
20 concentration of GDP in each well was 0.1 μ M GDP); each well comprising a candidate compound, has a final volume of 200 μ l consisting of 100 μ l GDP Buffer (final concentration, 0.1 μ M GDP), 50 μ l Membrane Protein in Binding Buffer, and 50 μ l [35 S]GTP γ S (0.6 nM) in Binding Buffer (2.5 μ l [35 S]GTP γ S per 10ml Binding Buffer).

b. Procedure

Candidate compounds will be preferably screened using a 96-well plate format (these can be frozen at -80°C). Membrane Protein (or membranes with expression vector excluding the GPCR Fusion Protein, as control), will be homogenized briefly until in suspension. Protein concentration will then be determined using, for example, the Bradford Protein Assay set forth above. Membrane Protein (and controls) will then be diluted to 0.25mg/ml in Binding Buffer (final assay concentration, 12.5µg/well). Thereafter, 100 µl GDP Buffer is added to each well of a Wallac Scintistrip™ (Wallac). A 5µl pin-tool will then be used to transfer 5 µl of a candidate compound into such well (*i.e.*, 5µl in total assay volume of 200 µl is a 1:40 ratio such that the final screening concentration of the candidate compound is 10µM). Again, to avoid contamination, after each transfer step the pin tool is rinsed in three reservoirs comprising water (1X), ethanol (1X) and water (2X) – excess liquid is shaken from the tool after each rinse and the tool is dried with paper and Kim wipes. Thereafter, 50 µl of Membrane Protein will be added to each well (a control well comprising membranes without the GPCR Fusion Protein was also utilized), and pre-incubated for 5-10 minutes at room temperature. Thereafter, 50 µl of [³⁵S]GTPγS (0.6 nM) in Binding Buffer will be added to each well, followed by incubation on a shaker for 60 minutes at room temperature (again, in this example, plates were covered with foil). The assay will be stopped by spinning the plates at 4000 RPM for 15 minutes at 22°C. The plates will then be aspirated with an 8 channel manifold and sealed with plate covers. The plates will then be read on a Wallac 1450 using setting “Prot. #37” (as per manufacturer's instructions).

B. Cyclic AMP Assay

Another assay approach to directly identify candidate compound will be accomplished utilizing a cyclase-based assay. In addition to direct identification, this assay

approach can be utilized as an independent approach to provide confirmation of the results from the [^{35}S]GTP γ S approach as set forth above.

A modified Flash PlateTM Adenylyl Cyclase kit (New England Nuclear; Cat. No. SMP004A) will be preferably utilized for direct identification of candidate compounds as
5 inverse agonists and agonists to GPCRs in accordance with the following protocol.

Transfected cells will be harvested approximately three days after transfection. Membranes will be prepared by homogenization of suspended cells in buffer containing 20mM HEPES, pH 7.4 and 10mM MgCl₂. Homogenization will be performed on ice using a Brinkman PolytronTM for approximately 10 seconds. The resulting homogenate will be
10 centrifuged at 49,000 X g for 15 minutes at 4°C. The resulting pellet will then be resuspended in buffer containing 20mM HEPES, pH 7.4 and 0.1 mM EDTA, homogenized for 10 seconds, followed by centrifugation at 49,000 X g for 15 minutes at 4°C. The resulting pellet will then be stored at -80°C until utilized. On the day of direct identification screening, the membrane pellet will slowly be thawed at room temperature, resuspended in
15 buffer containing 20mM HEPES, pH 7.4 and 10mM MgCl₂, to yield a final protein concentration of 0.60mg/ml (the resuspended membranes will be placed on ice until use).

cAMP standards and Detection Buffer (comprising 2 μCi of tracer [^{125}I] cAMP (100 μl) to 11 ml Detection Buffer) will be prepared and maintained in accordance with the manufacturer's instructions. Assay Buffer will be prepared fresh for screening and contain
20 20mM HEPES, pH 7.4, 10mM MgCl₂, 20mM phosphocreatine (Sigma), 0.1 units/ml creatine phosphokinase (Sigma), 50 μM GTP (Sigma), and 0.2 mM ATP (Sigma); Assay Buffer will be stored on ice until utilized.

Candidate compounds identified as per above (if frozen, thawed at room temperature) will be added, preferably, to 96-well plate wells (3 μl /well; 12 μM final assay

concentration), together with 40 μ l Membrane Protein (30 μ g/well) and 50 μ l of Assay Buffer. This admixture will be incubated for 30 minutes at room temperature, with gentle shaking.

Following the incubation, 100 μ l of Detection Buffer will be added to each well,
5 followed by incubation for 2-24 hours. Plates will then be counted in a Wallac MicroBeta™ plate reader using "Prot. #31" (as per manufacturer instructions).

C. Melanophore Screening Assay

A method for identifying candidate agonists or inverse agonists for a GPCR can be preformed by introducing tests cells of a pigment cell line capable of dispersing or
10 aggregating their pigment in response to a specific stimulus and expressing an exogenous clone coding for the GCPR. A stimulant, e.g., light, sets an initial state of pigment disposition wherein the pigment is aggregated within the test cells if activation of the GPCR induces pigment dispersion. However, stimulating the cell with a stimulant to set an initial state of pigment disposition wherein the pigment is dispersed if activation of the GPCR
15 induces pigment aggregation. The tests cells are then contacted with chemical compounds, and it is determined whether the pigment disposition in the cells changed from the initial state of pigment disposition. Dispersion of pigments cells due to the candidate compound coupling to the GPCR will appear dark on a petri dish, while aggregation of pigments cells will appear light.

20 Materials and methods will be followed according to the disclosure of U.S. Patent Number 5,462,856 and U.S. Patent Number 6,051,386, each of which are incorporated by reference in its entirety.

Although a variety of expression vectors are available to those in the art, for purposes of utilization for both the endogenous and non-endogenous human GPCRs, in

some embodiments it is preferred that the vector utilized be pCMV. This vector was deposited with the American Type Culture Collection (ATCC) on October 13, 1998 (10801 University Blvd., Manassas, VA 20110-2209 USA) under the provisions of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purpose
5 of Patent Procedure. The DNA was tested by the ATCC and determined to be viable. The ATCC has assigned the following deposit number to pCMV: ATCC #203351.

References cited throughout this patent document, including co-pending and related patent applications, unless otherwise indicated, are fully incorporated herein by reference. Modifications and extension of the disclosed inventions that are within the purview of the
10 skilled artisan are encompassed within the above disclosure and the claims that follow.

CLAIMS

What is claimed is:

1. A G protein-coupled receptor encoded by an amino acid sequence of
5 SEQ.ID.NO.:2.
2. A non-endogenous, constitutively activated version of the G protein-coupled
receptor of claim 1.
3. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:1.
4. A host cell comprising the plasmid of claim 3.
- 10 5. A G protein-coupled receptor encoded by an amino acid sequence of
SEQ.ID.NO.:4.
6. A non-endogenous, constitutively activated version of the G protein-coupled
receptor of claim 5.
7. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:3.
- 15 8. A host cell comprising the plasmid of claim 7.
9. A G protein-coupled receptor encoded by an amino acid sequence of
SEQ.ID.NO.:6.
10. A non-endogenous, constitutively activated version of the G protein-coupled
receptor of claim 9.
- 20 11. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:5.
12. A host cell comprising the plasmid of claim 11.
13. A G protein-coupled receptor encoded by an amino acid sequence of
SEQ.ID.NO.:8.

14. A non-endogenous, constitutively activated version of the G protein-coupled receptor of claim 13.
15. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:7.
16. A host cell comprising the plasmid of claim 15.
- 5 17. A G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:10.
18. A non-endogenous, constitutively activated version of the G protein-coupled receptor of claim 17 .
19. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:9.
- 10 20. A host cell comprising the plasmid of claim 19.
21. A G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:12.
22. A non-endogenous, constitutively activated version of the G protein-coupled receptor of claim 21.
- 15 23. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:11.
24. A host cell comprising the plasmid of claim 23.
25. A G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:14.
26. A non-endogenous, constitutively activated version of the G protein-coupled receptor of claim 25.
- 20 27. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:13.
28. A host cell comprising the plasmid of claim 27.
29. A G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:16.

30. A non-endogenous, constitutively activated version of the G protein-coupled receptor of claim 29.
31. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:15.
32. A host cell comprising the plasmid of claim 31.
- 5 33. A G protein-coupled receptor encoded by an amino acid sequence of SEQ.ID.NO.:18.
34. A non-endogenous, constitutively activated version of the G protein-coupled receptor of claim 33.
35. A plasmid comprising a vector and the cDNA of SEQ.ID.NO.:17.
- 10 36. A host cell comprising the plasmid of claim 35.

Figure 1
Cell-Based cAMP Assay

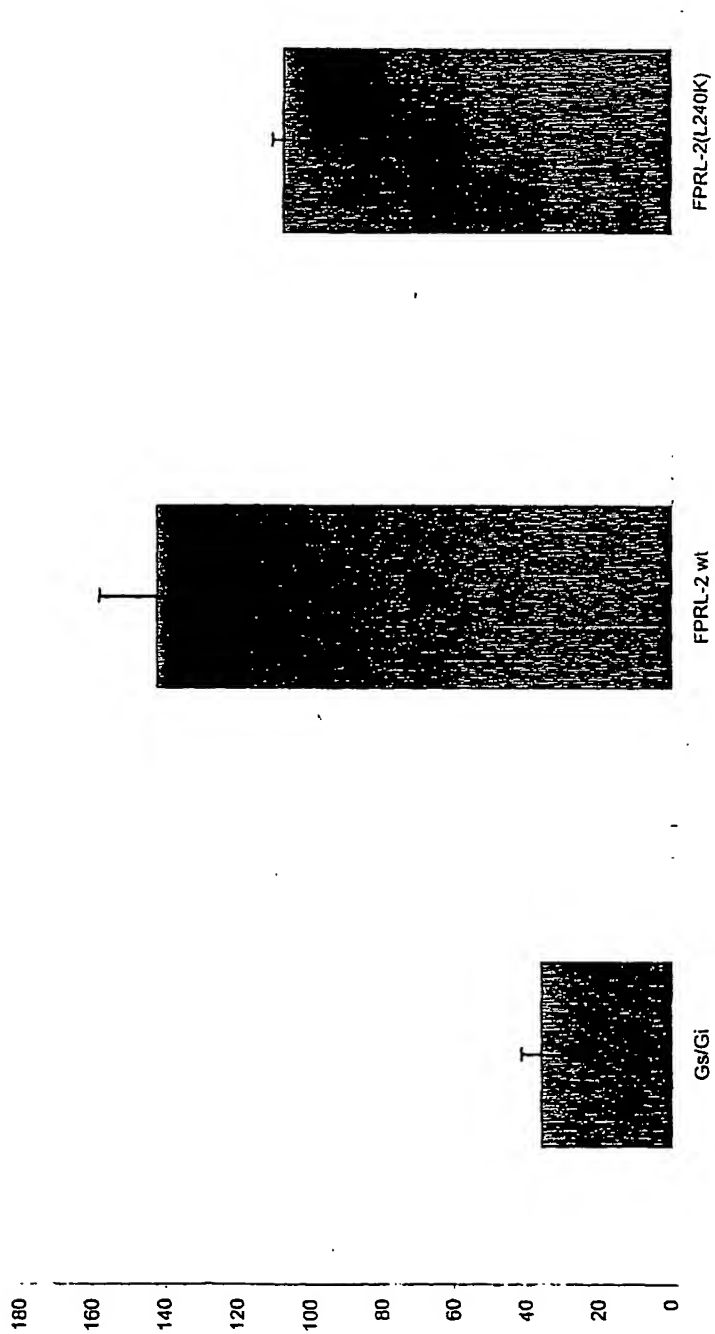
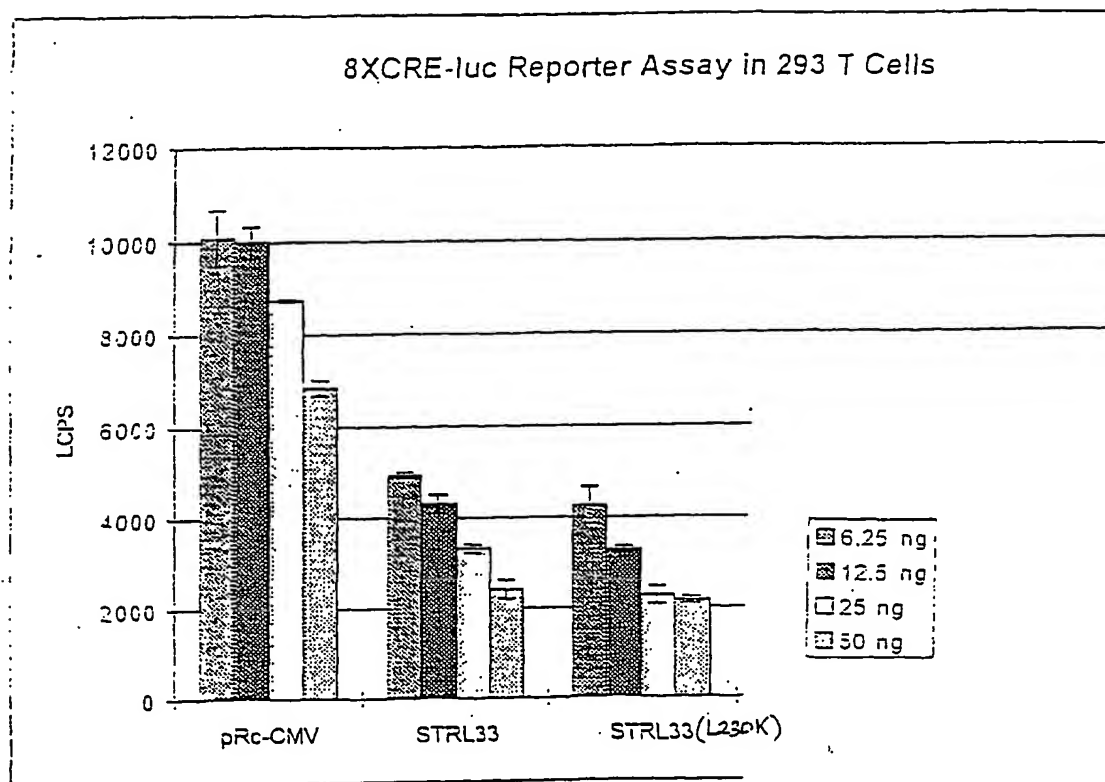


Figure 2



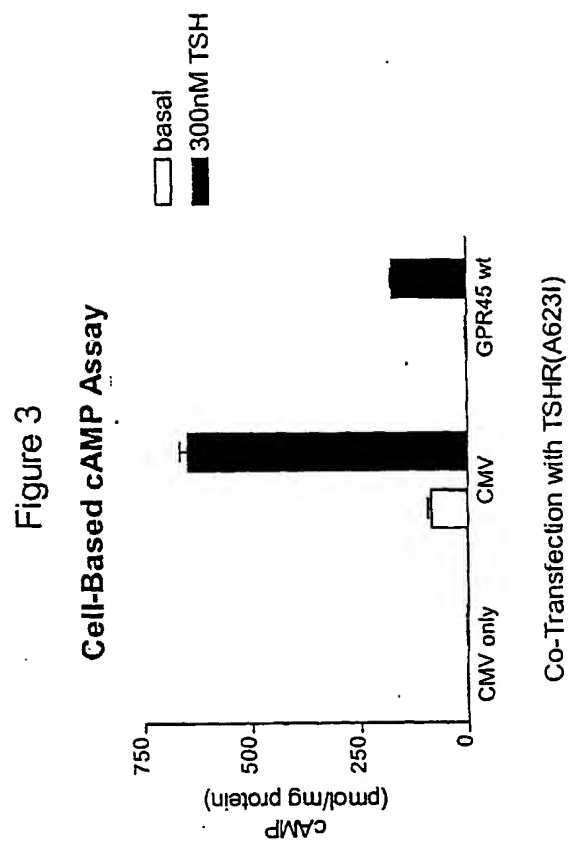
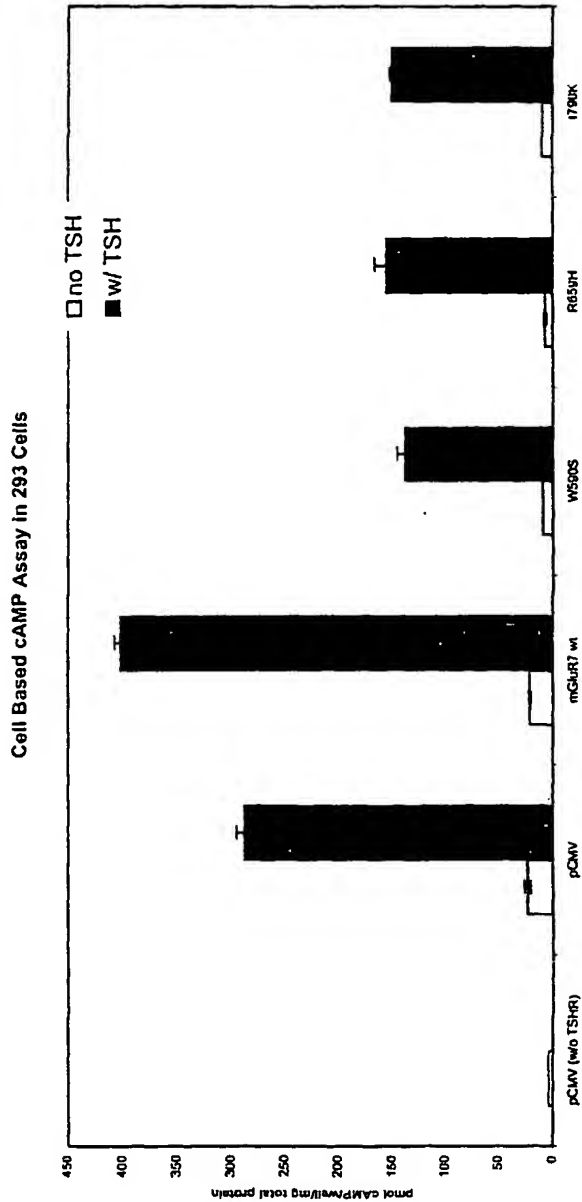


Figure 4



Co-Transfection with TSHR(A623I)

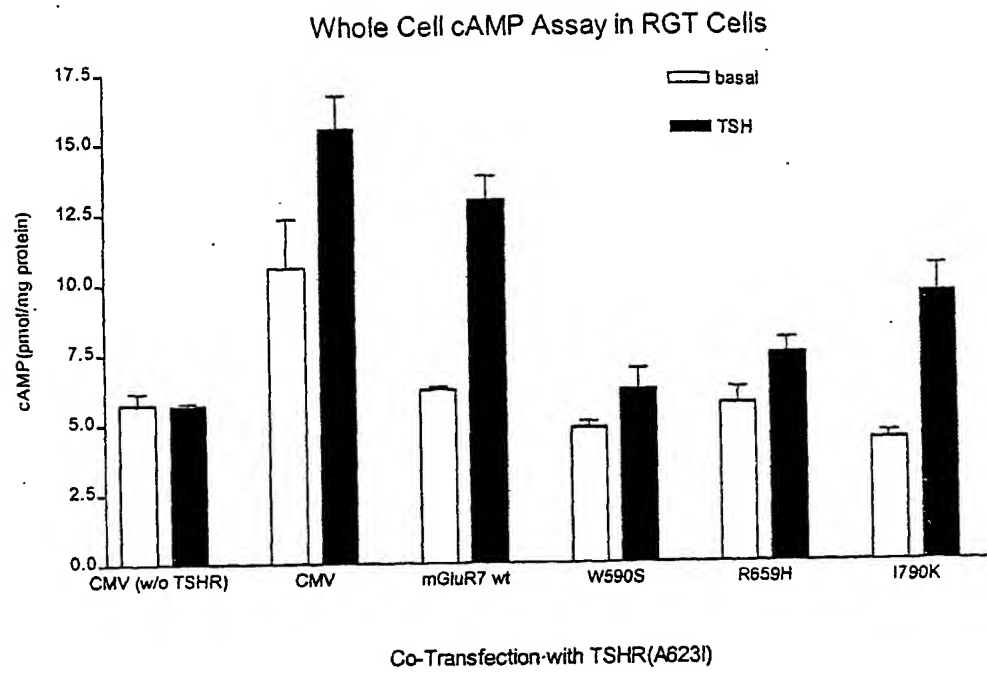


Figure 5

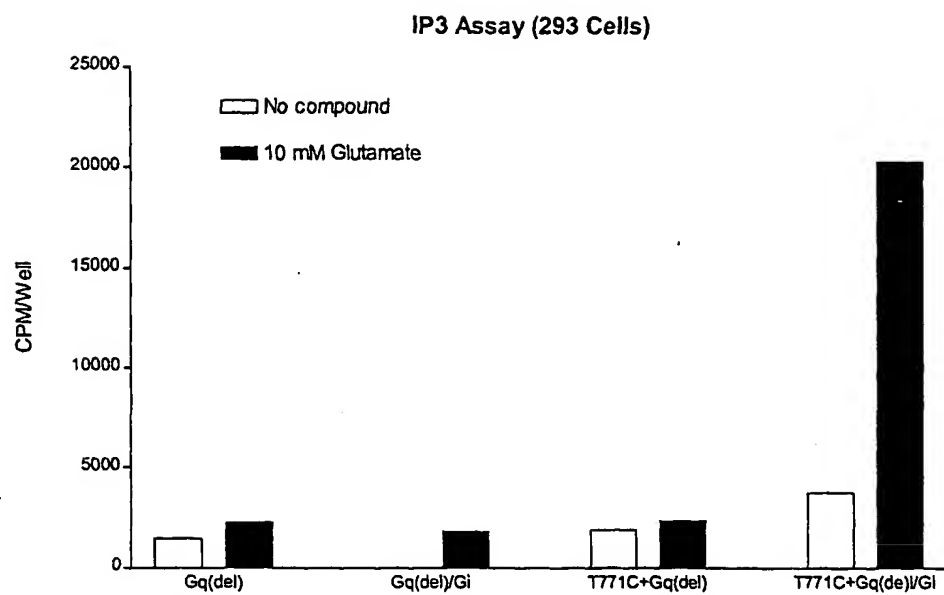
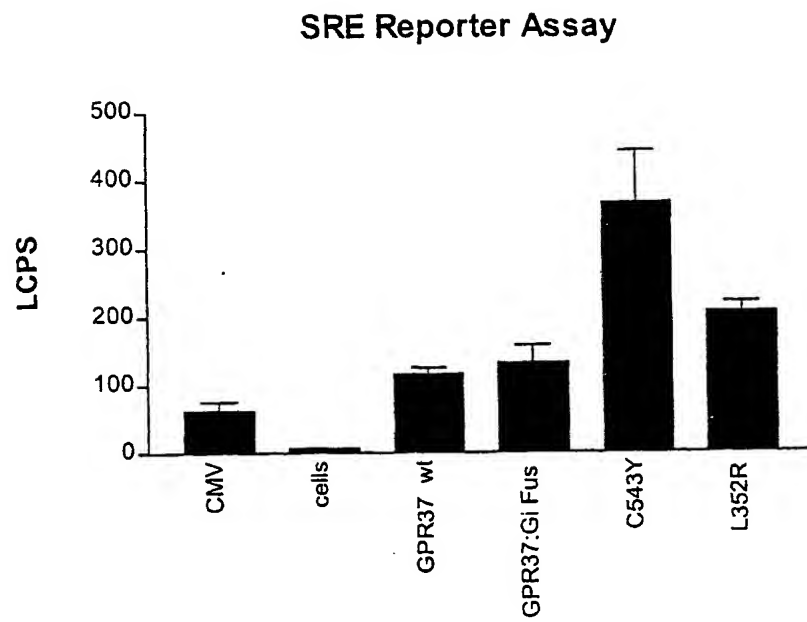
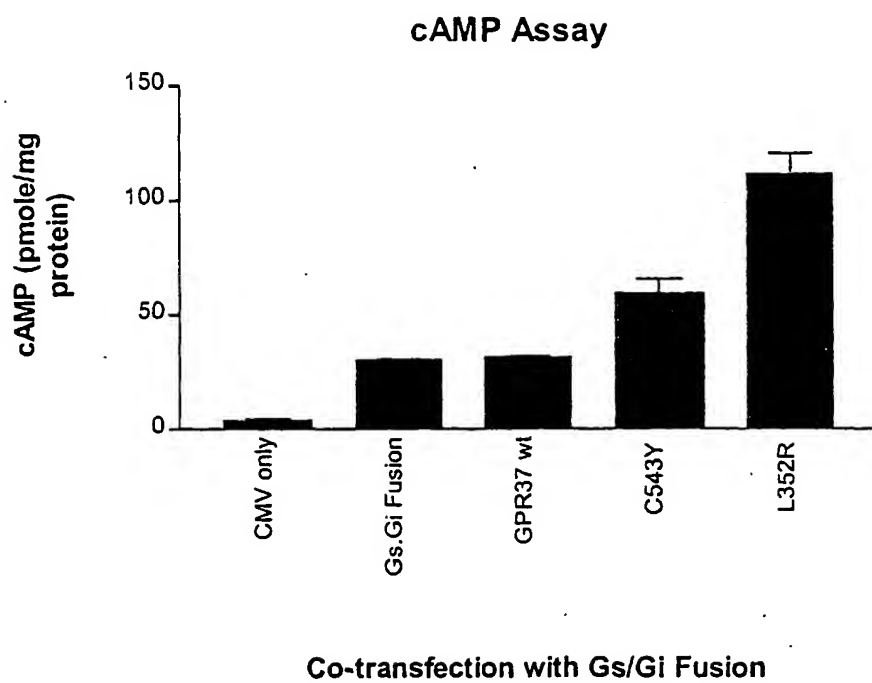


Figure 6

**Figure 7**

**Figure 8**

Northern Analysis of GPCR GPR37
expression in forskolin treated Rat
Schwann cells

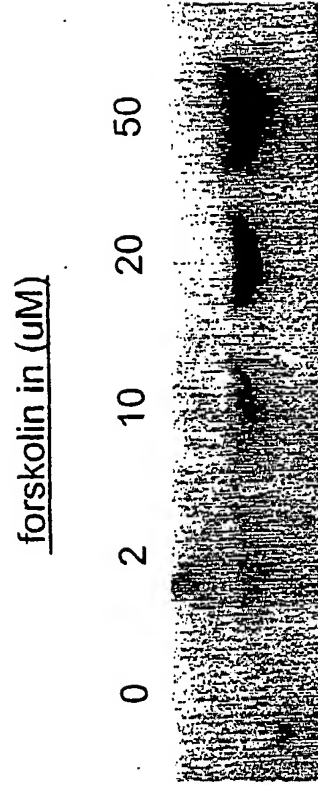


Figure 9

Northern Analysis of GPCR GPR37
Expression in Crushed Rat Sciatic Nerve

Days post-crush

0 1 3 7 10 13



Figure 10

11/21

IP3 Assay

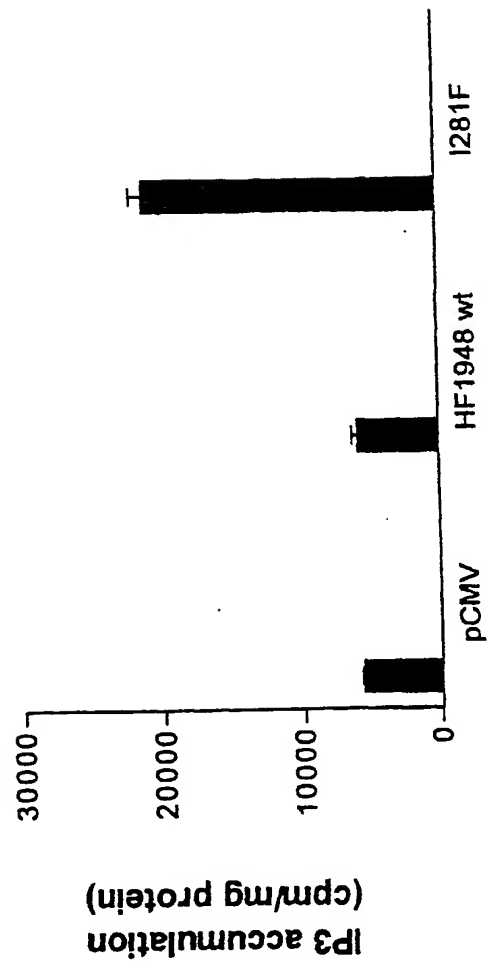


Figure 11

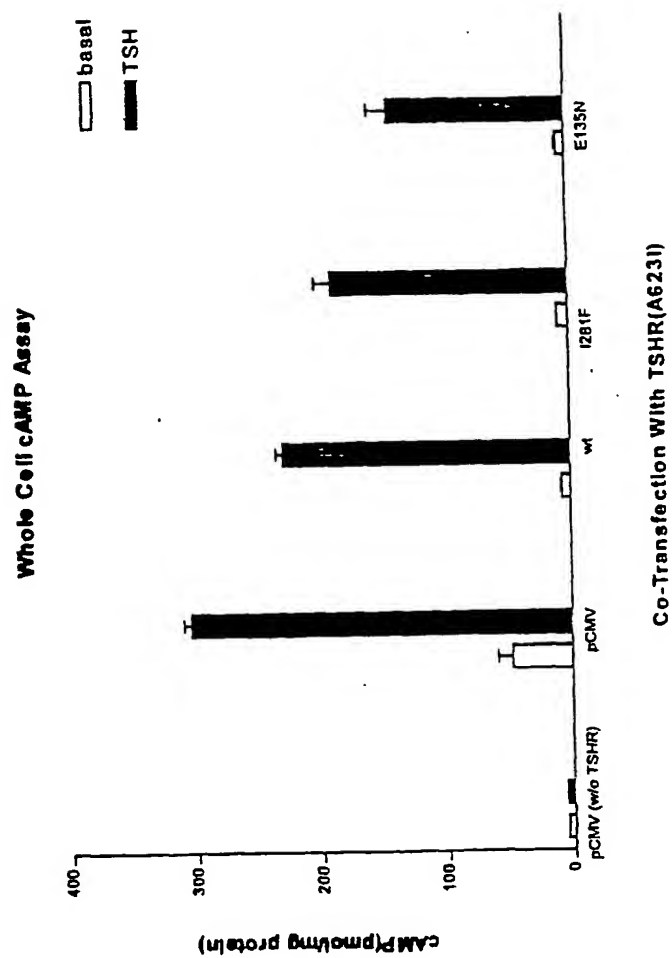
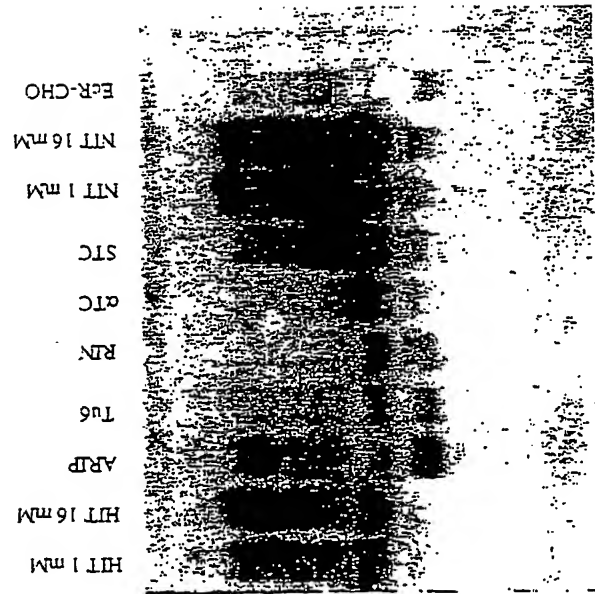


Figure 12

Figure 13

Cell-specific expression of GPR66 variants in pancreatic cell lines



E2F-Luc activation by GPR35 in 293A cells

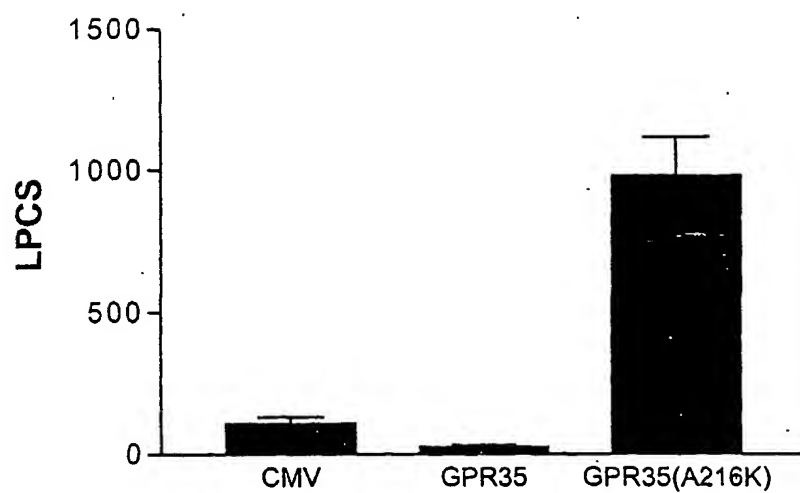
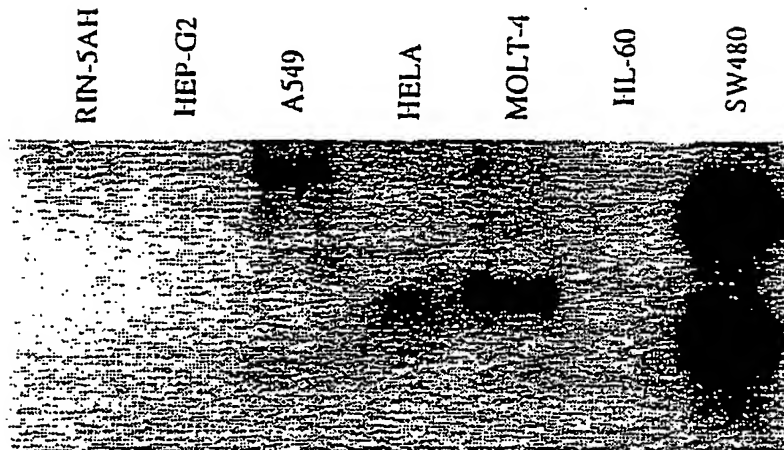
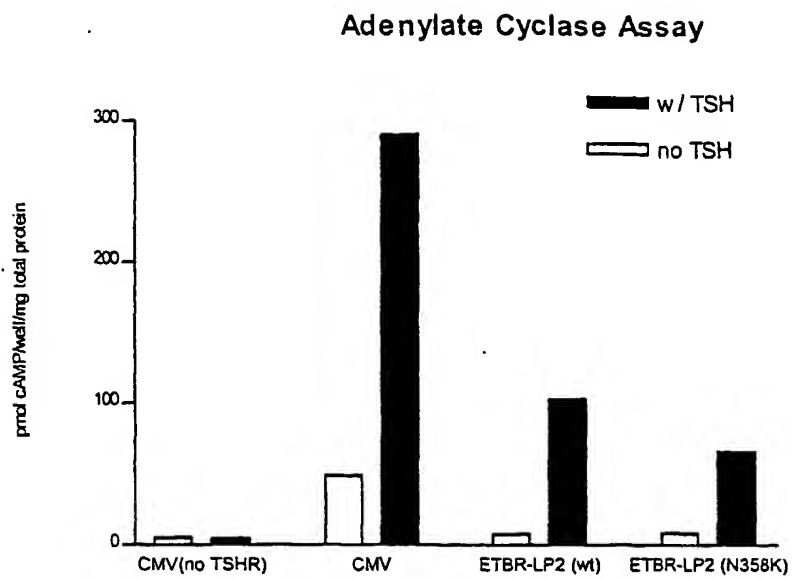


Figure 14

Figure 15

**Expression of GPR35
in colorectal cancer cells**





Co-Transfection with TSHR(A623I)

Figure 16

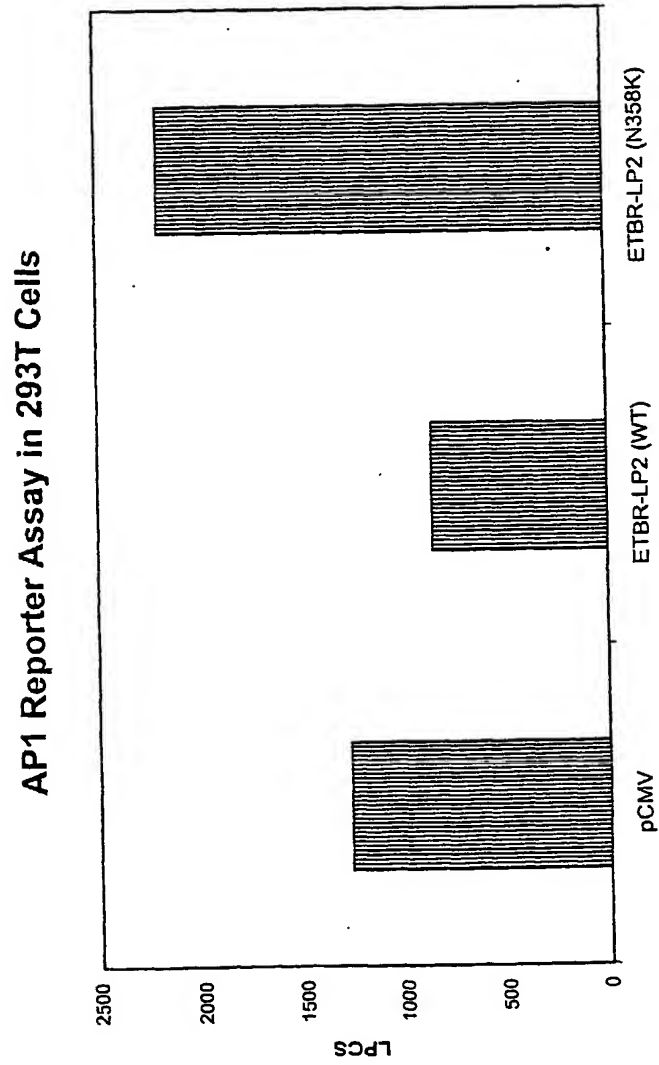


Figure 17

**Northern Analysis of ETBR-LP2 in
Forskolin Treated Rat Schwann Cells**

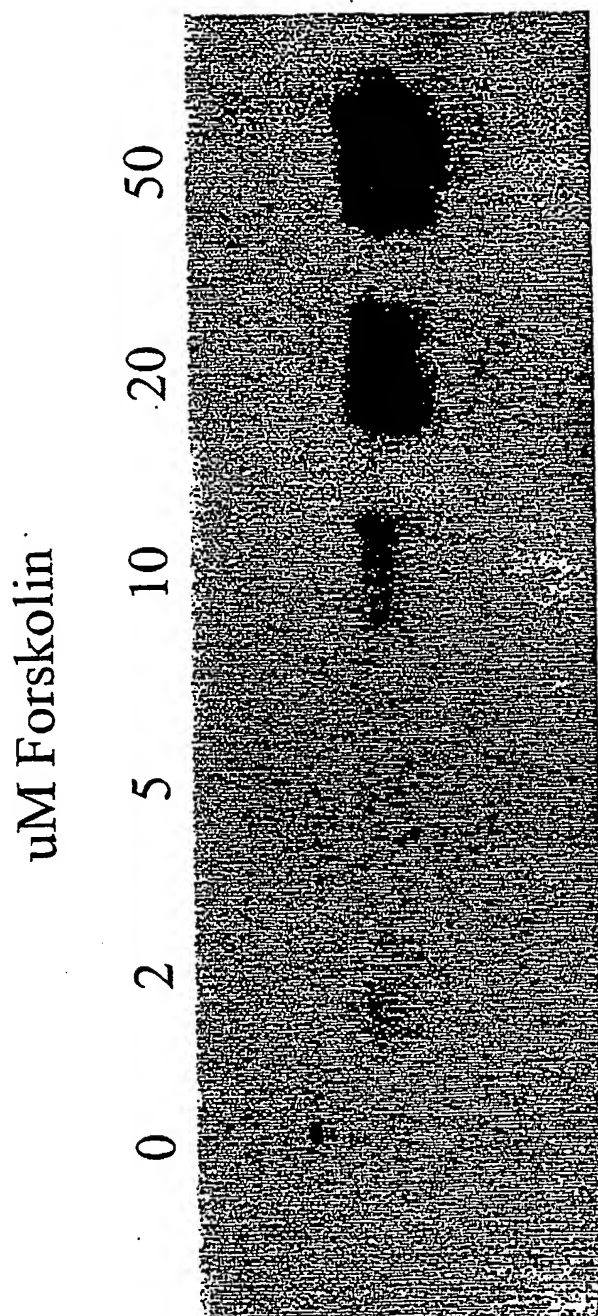


Figure 18

Northern Analysis of ETBR-LP2 Expression in
Crushed Rat Sciatic Nerve

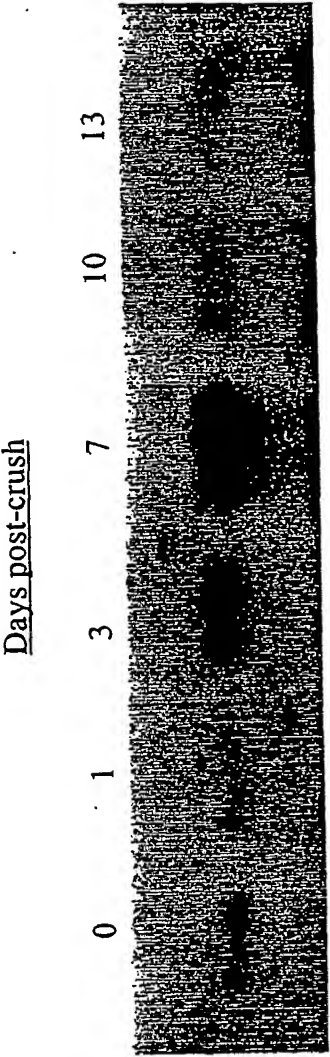


Figure 19

Figure 20A

		M R A L G A L A A S L A V L L A V G L L K V S G G A A L G V G P A S R N E T C L	Majority
		10 20 30 40	
1		M R W L W P L A V I S L A V I L A V G L S R V S G G A P L H L G - - - - -	HETERLP2p
1		M R A P G A L L A R M S R L L L L L L L K V S A S S A L G V A P A S R N E T C L	HGPR37p
		G E S C A P T V I O R R G R D A W G P G N S A R D V L R A R A E T E E Q G A A F	Majority
		50 60 70 80	
32		- - - - - R H R A E T Q E Q Q S - -	HETERLP2p
41		G E S C A P T V I O R R G R D A W G P G N S A R D V L R A R A P R E E Q G A A F	HGPR37p
		L A G P S W D L P A A P G R D P A A G R G A E A S A A G P P G P P T R P P G P W	Majority
		90 100 110 120	
43		- - - - -	HETERLP2p
61		L A G P S W D L P A A P G R D P A A G R G A E A S A A G P P G P P T R P P G P W	HGPR37p
		R W K G A R G T E P S E T L G R G N P T A L Q L F L O I S D E E A K G V Q G A G	Majority
		130 140 150 160	
43		- - R S K R G T E - - - - - D E E A K G V Q - - -	HETERLP2p
121		R W K G A R G Q E P S E T L G R G N P T A L Q L F L O I S E E E K G P R G A G	HGPR37p
		I S G R S Q E Q S V Q T V P G A S A L F Y R P I H A G G L O G S H H K P L V A T	Majority
		170 180 190 200	
58		- - - - - Q Y V F E E W A E Y P R P I H P A G L Q P T - - K P L V A T	HETERLP2p
161		I S G R S Q E Q S V K T V P G A S D L F Y W F R R A G K L O G S H H K P L S K T	HGPR37p
		A N G L A G D G G W T I A L P G S G L A L N G S L G G G I H E P G G P R R G N S	Majority
		210 220 230 240	
86		S P N P D K D G G - - - - T F D S G Q E L R G N L T G A - - - P G - - - - -	HETERLP2p
201		A N G L A G H E G W T I A L P G R A L A O N G S L G E G I H E P G G P R R G N S	HGPR37p
		T N Q R V Q L Q N P L Y P V T E S S Y G A Y A V M L L A V V V F G V G I V G N L	Majority
		250 260 270 280	
112		- - Q R L Q I Q N P L Y P V T E S S Y S A Y A I M L L A L V V F A V G I V G N L	HETERLP2p
241		T N R R V R L K N E F Y E L T Q E S Y G A Y A V M C L S V V I F G T G I I G N L	HGPR37p
		A V M C I V W H S Y Y L K S A S N S L L A S L A L W D F L V L F F C L P L V I F	Majority
		290 300 310 320	
150		S V M C I V W H S Y Y L K S A W N S I L A S L A L W D F L V L F F C L P I V I F	HETERLP2p
281		A V M C I V C H N Y Y M R S I S N S L L A N L A F W D F L I I F F C L P L V I F	HGPR37p
		N E L T K Q R L L G D V S C K A V P F I E V A S L G V T T F S L C A L G I D R F	Majority
		330 340 350 360	
190		N E I T K Q R L L G D V S C R A V P F E V S S L G V T T F S L C A L G I D R F	HETERLP2p
321		H E L T K K W L L E D F S C K I V E Y I E V A S L G V T T F T L C A L C I D R F	HGPR37p

Figure 20B

	HAATSVLMKVEMIENCSSILAKLAVIWVGALLLAVPEVVL																														Majority											
	370										380										390										400											
230	H	V	A	T	S	T	L	P	K	V	R	P	I	E	R	C	Q	S	I	L	A	K	L	A	V	I	W	V	G	S	M	T	L	A	V	P	E	L	L	L		HETERLP2p
361	R	A	A	T	N	V	Q	M	Y	Y	E	M	I	E	N	C	S	S	T	T	A	K	L	A	V	I	W	V	G	A	L	L	L	A	L	P	E	V	V	L		HGPR37p
	RQLAQEDAGFSGRGTADSCI IKISASLPDSL YVLALTYDS																														Majority											
	410										420										430										440											
270	W	Q	L	A	O	E	P	A	-	-	P	T	M	G	T	L	D	S	C	I	M	K	P	S	A	S	L	P	E	S	L	Y	S	L	V	M	T	Y	Q	N		HETERLP2p
401	R	O	L	S	K	E	D	L	G	F	S	G	R	A	P	A	E	R	C	I	K	I	S	P	D	L	P	D	T	I	Y	V	L	A	L	T	Y	D	S		HGPR37p	
	ARLWWYFGCYFCLPILFTVTC SLVTARKV RGAPGRESACT																														Majority											
	450										460										470										480											
308	A	R	M	W	W	Y	F	G	C	Y	F	C	L	P	I	L	F	T	V	T	C	Q	L	V	T	-	W	R	V	R	G	P	P	G	R	K	S	E	C	-		HETERLP2p
441	A	R	L	W	W	Y	F	G	C	Y	F	C	L	P	T	L	F	T	I	T	C	S	L	V	T	A	R	K	I	R	K	A	-	-	-	E	K	A	C	T		HGPR37p
	RGSKEHIOLESQ LNSTVVG LTVVYGFCILPENVCNIVVAY																														Majority											
	490										500										510										520											
346	R	A	S	K	H	E	-	Q	C	E	S	Q	L	N	S	T	V	V	G	L	T	V	V	Y	A	F	C	T	L	P	E	N	V	C	N	I	V	V	A	Y		HETERLP2p
478	R	G	N	K	R	Q	I	O	E	S	O	M	N	C	T	V	V	A	L	T	I	L	Y	G	F	C	I	I	P	E	N	I	C	N	I	V	T	A	Y		HGPR37p	
	LATGV SQQTLDLLGLISQFL LFFKGA VTPVLLCLCKPLG																														Majority											
	530										540										550										560											
385	L	S	T	E	L	T	R	Q	T	L	D	L	L	G	L	I	N	Q	F	S	T	F	F	K	G	A	I	T	P	V	L	L	L	C	I	C	R	P	L	G		HETERLP2p
518	M	A	T	G	V	S	O	O	T	M	D	L	L	N	I	I	S	O	F	L	L	F	F	K	S	C	V	T	P	V	L	L	F	C	L	C	K	P	F	S		HGPR37p
	QAFLDCCCCCCEECGGASSAVAADGSDNELTTEVSLSIF																														Majority											
	570										580										590										600											
425	Q	A	F	L	D	C	C	C	C	C	C	C	E	E	C	G	G	A	S	E	A	S	A	A	N	G	S	D	N	K	L	K	T	E	V	S	S	S	I	Y		HETERLP2p
558	R	A	F	M	E	C	C	C	C	C	C	-	E	E	C	I	Q	K	S	S	T	V	T	S	D	D	N	D	N	E	Y	T	T	E	L	E	L	S	P	F		HGPR37p
	STIRRESSTLASVGTHC																														Majority											
	610																																									
465	F	H	K	P	R	E	S	P	P	L	L	P	L	G	T	P	C																			HETERLP2p						
597	S	T	I	R	R	E	M	S	T	F	A	S	V	G	T	H	C																			HGPR37p						

Decoration 'Decoration #1': Box residues that match the Consensus exactly.

Decoration 'Decoration #2': Box residues that match the Consensus exactly.

SEQUENCE LISTING

<110> Arena Pharmaceuticals, Inc.

<120> Endogenous And Non-Endogenous, Constitutively Activated G Protein-Coupled Receptors

<130> AREN-0321

<160> 102

<170> PatentIn version 3.1

<210> 1

<211> 1062

<212> DNA

<213> Homo sapiens

<400> 1

atggaaacca acttctccat tcctctgaat gaaactgagg aggtgctccc tgagcctgct	60
ggccacaccg ttctgtggat cttctcattg ctagtccacg gagtcacctt tgtcttcggg	120
gtcctgggca atgggcttgt gatctgggtg gctggattcc ggatgacacg cacagtcaac	180
accatctgtt acctgaacct ggccctagct gacttctctt tcagtgccat cctaccattc	240
cgaatgggtct cagtcgccat gagagaaaaa tggccttttg gtcattcct atgtaagtta	300
gttcatgtta tgatagacat caacctgttt gtcagtgtct acctgatcac catcattgct	360
ctggaccgct gtatttgtgt cctgcatcca gcctgggccc agaaccatcg caccatgagt	420
ctggccaaga gggatgatgac gggactctgg attttcacca tagtccttac cttaccaa	480
ttcatcttct ggactacaat aagtactacg aatggggaca catactgtat tttcaacttt	540
gcattctggg gtgacactgc ttagagagg ttgaacgtgt tcattaccat ggccaaggtc	600
tttctgatcc tccacttcat tattggcttc agcgtgccta tgtccatcat cacagtctgc	660
tatgggatca tcgctgcaa aattcacaga aaccacatga ttaaattccag ccgtccctta	720
cgtgtcttcg ctgctgtggt ggcttcttct ttcattctgtt gggtccctta tgaactaatt	780
ggcattctaa tggcagtctg gctcaaagag atgttggtta atggcaaata caaaatcatt	840
cttgtctga ttaacccaac aagctccttg gcctttttta acagctgcct caacccaatt	900
ctctacgtct ttatgggtcg taacttccaa gaaagactga ttcgctcttt gccactagt	960
ttggagagg ccctgactga ggtccctgac tcagcccaga ccagcaacac agacaccact	1020
tctgcttcac ctctgagga gacggagtta caagcaatgt ga	1062

<210> 2
 <211> 353
 <212> PRT
 <213> Homo sapiens

<400> 2

Met Glu Thr Asn Phe Ser Ile Pro Leu Asn Glu Thr Glu Glu Val Leu
 1 5 10 15

Pro Glu Pro Ala Gly His Thr Val Leu Trp Ile Phe Ser Leu Leu Val
 20 25 30

His Gly Val Thr Phe Val Phe Gly Val Leu Gly Asn Gly Leu Val Ile
 35 40 45

Trp Val Ala Gly Phe Arg Met Thr Arg Thr Val Asn Thr Ile Cys Tyr
 50 55 60

Leu Asn Leu Ala Leu Ala Asp Phe Ser Phe Ser Ala Ile Leu Pro Phe
 65 70 75 80

Arg Met Val Ser Val Ala Met Arg Glu Lys Trp Pro Phe Gly Ser Phe
 85 90 95

Leu Cys Lys Leu Val His Val Met Ile Asp Ile Asn Leu Phe Val Ser
 100 105 110

Val Tyr Leu Ile Thr Ile Ile Ala Leu Asp Arg Cys Ile Cys Val Leu
 115 120 125

His Pro Ala Trp Ala Gln Asn His Arg Thr Met Ser Leu Ala Lys Arg
 130 135 140

Val Met Thr Gly Leu Trp Ile Phe Thr Ile Val Leu Thr Leu Pro Asn
 145 150 155 160

Phe Ile Phe Trp Thr Thr Ile Ser Thr Thr Asn Gly Asp Thr Tyr Cys
 165 170 175

Ile Phe Asn Phe Ala Phe Trp Gly Asp Thr Ala Val Glu Arg Leu Asn
 180 185 190

Val Phe Ile Thr Met Ala Lys Val Phe Leu Ile Leu His Phe Ile Ile
 195 200 205

Gly Phe Ser Val Pro Met Ser Ile Ile Thr Val Cys Tyr Gly Ile Ile
 210 215 220

Ala Ala Lys Ile His Arg Asn His Met Ile Lys Ser Ser Arg Pro Leu
 225 230 235 240

Arg Val Phe Ala Ala Val Val Ala Ser Phe Phe Ile Cys Trp Phe Pro
 245 250 255

Tyr Glu Leu Ile Gly Ile Leu Met Ala Val Trp Leu Lys Glu Met Leu
 260 265 270

Leu Asn Gly Lys Tyr Lys Ile Ile Leu Val Leu Ile Asn Pro Thr Ser
 275 280 285

Ser Leu Ala Phe Phe Asn Ser Cys Leu Asn Pro Ile Leu Tyr Val Phe
 290 295 300

Met Gly Arg Asn Phe Gln Glu Arg Leu Ile Arg Ser Leu Pro Thr Ser
 305 310 315 320

Leu Glu Arg Ala Leu Thr Glu Val Pro Asp Ser Ala Gln Thr Ser Asn
 325 330 335

Thr Asp Thr Thr Ser Ala Ser Pro Pro Glu Glu Thr Glu Leu Gln Ala
 340 345 350

Met

<210> 3
 <211> 1029
 <212> DNA
 <213> Homo sapiens

<400> 3
 atggcagagc atgattacca tgaagactat gggttcagca gtttcaatga cagcagccag 60
 gaggagcatc aagccttctt gcagttcagc aagggtctttc tgccctgcat gtacctggtg 120
 gtgtttgtct gtggtctggt ggggaactct ctggtgctgg tcatatccat cttctaccat 180
 aagttgcaga gcctgacgga tgtgttcctg gtgaacctac ccctggctga cctggtgttt 240
 gtctgcactc tgcccttctg ggcctatgca ggcacccatg aatgggtgtt tggccaggtc 300

```

atgtgcaaaa gcctactggg catctacact attaacttct acacgtccat gctcatcctc 360
acctgcatca ctgtggatcg ttctattgta gtggttaagg ccaccaaggc ctacaaccag 420
caagccaaga ggatgacctg gggcaaggtc accagcttgc tcctctgggt gatatccctg 480
ctgggttctc tgcctccaaat tatctatggc aatgtcttta atctcgacaa gctcatatgt 540
ggttaccatg acgaggcaat ttccactgtg gttcttgcca ccagatgac actgggggttc 600
ttcttgccac tgcctcaccat gattgtctgc tattcagtca taatcaaaac actgcttcat 660
gctggaggct tccagaagca cagatctcta aagatcatct tcctgggtgat ggctgtgttc 720
ctgctgaccc agatgccctt caacctcatg aagttcatcc gcagcacaca ctgggaatac 780
tatgccatga ccagctttca ctacaccatc atggtgacag aggccatcgc atacctgagg 840
gcctgcctta accctgtgct ctatgccttt gtcagcctga agtttcgaaa gaacttctgg 900
aaacttgtga aggacattgg ttgcctccct taccttgggg tctcacatca atggaaatct 960
tctgaggaca attccaagac tttttctgcc tcccacaatg tggaggccac cagcatgttc 1020
cagttatag 1029

```

```

<210> 4
<211> 342
<212> PRT
<213> Homo sapiens

```

```

<400> 4

```

```

Met Ala Glu His Asp Tyr His Glu Asp Tyr Gly Phe Ser Ser Phe Asn
1          5          10          15

```

```

Asp Ser Ser Gln Glu Glu His Gln Ala Phe Leu Gln Phe Ser Lys Val
20          25          30

```

```

Phe Leu Pro Cys Met Tyr Leu Val Val Phe Val Cys Gly Leu Val Gly
35          40          45

```

```

Asn Ser Leu Val Leu Val Ile Ser Ile Phe Tyr His Lys Leu Gln Ser
50          55          60

```

```

Leu Thr Asp Val Phe Leu Val Asn Leu Pro Leu Ala Asp Leu Val Phe
65          70          75          80

```

```

Val Cys Thr Leu Pro Phe Trp Ala Tyr Ala Gly Ile His Glu Trp Val
85          90          95

```


Phe Gly Gln Val Met Cys Lys Ser Leu Leu Gly Ile Tyr Thr Ile Asn
 100 105 110

Phe Tyr Thr Ser Met Leu Ile Leu Thr Cys Ile Thr Val Asp Arg Phe
 115 120 125

Ile Val Val Val Lys Ala Thr Lys Ala Tyr Asn Gln Gln Ala Lys Arg
 130 135 140

Met Thr Trp Gly Lys Val Thr Ser Leu Leu Ile Trp Val Ile Ser Leu
 145 150 155 160

Leu Val Ser Leu Pro Gln Ile Ile Tyr Gly Asn Val Phe Asn Leu Asp
 165 170 175

Lys Leu Ile Cys Gly Tyr His Asp Glu Ala Ile Ser Thr Val Val Leu
 180 185 190

Ala Thr Gln Met Thr Leu Gly Phe Phe Leu Pro Leu Leu Thr Met Ile
 195 200 205

Val Cys Tyr Ser Val Ile Ile Lys Thr Leu Leu His Ala Gly Gly Phe
 210 215 220

Gln Lys His Arg Ser Leu Lys Ile Ile Phe Leu Val Met Ala Val Phe
 225 230 235 240

Leu Leu Thr Gln Met Pro Phe Asn Leu Met Lys Phe Ile Arg Ser Thr
 245 250 255

His Trp Glu Tyr Tyr Ala Met Thr Ser Phe His Tyr Thr Ile Met Val
 260 265 270

Thr Glu Ala Ile Ala Tyr Leu Arg Ala Cys Leu Asn Pro Val Leu Tyr
 275 280 285

Ala Phe Val Ser Leu Lys Phe Arg Lys Asn Phe Trp Lys Leu Val Lys
 290 295 300

Asp Ile Gly Cys Leu Pro Tyr Leu Gly Val Ser His Gln Trp Lys Ser
 305 310 315 320

Ser Glu Asp Asn Ser Lys Thr Phe Ser Ala Ser His Asn Val Glu Ala
 325 330 335

Thr Ser Met Phe Gln Leu
 340

<210> 5
 <211> 1119
 <212> DNA
 <213> Homo sapiens

<400> 5
 atggcctgca acagcacgtc ccttgaggct tacacatacc tgctgctgaa caccagcaac 60
 gcctcagact cggggtccac ccagttgccc gcacccctca ggatctcctt ggccatagtg 120
 atgctgctga tgaccgtggt ggggttcctg ggcaacactg tggctctgcat catcgtgtac 180
 cagaggccgg ctatgcgctc ggccatcaac ctgctgctgg ccacctggc cttctccgac 240
 atcatgctgt ccctctgctg catgcccttc accgccgtca ccctcatcac cgtgcgctgg 300
 cactttgggg accacttctg ccgcctctca gccacgctct actgggtttt tgtcctggag 360
 ggcgtggcca tcctgctcat catcagcgtg gaccgcttc tcatcatcgt ccagcgccag 420
 gacaagctga acccgcgag ggccaagggt atcatcgcg tctcctgggt gctgtccttc 480
 tgcacgcgg ggcctcgt caccggctgg acgctgggtg aggtgccggc gcgggcccc 540
 cagtgcgtgc tgggtacac ggagctcccc gctgaccgag catacgtggt caccttggtg 600
 gtggcgtgt tcttcgcgcc ctttgcgctc atgctgtgag cctacatgtg catcctcaac 660
 acggtccgca agaacgccgt gcgcgtgcac aaccagtcgg acagcctgga cctgcggcag 720
 ctaccaggc cgggcctgcg gcgcctgcag cggcagcaac aggtcagcgt ggacttgagc 780
 ttcaagacca aggccttcac caccatcctg atcctcttcg tgggcttctc cctctgctgg 840
 ctgccccact ccgtctacag cctcctgtct gtgttttagc agcgctttta ctgcggttcc 900
 tccttctacg ccaccagcac ctgcgtcctg tgggtcagtt acctcaagtc cgtcttcaac 960
 cccatcgtct actgctggag aatcaaaaaa ttccgcgagg cctgcataga gttgctgccc 1020
 cagaccttc aaatcctccc caaagtgcct gagcggtacc gaaggagaat ccagccaagc 1080
 acagtatacg tgtgcaatga aaaccagtct gcggtttag 1119

<210> 6
 <211> 372
 <212> PRT

<213> Homo sapiens

<400> 6

Met Ala Cys Asn Ser Thr Ser Leu Glu Ala Tyr Thr Tyr Leu Leu Leu
 1 5 10 15

Asn Thr Ser Asn Ala Ser Asp Ser Gly Ser Thr Gln Leu Pro Ala Pro
 20 25 30

Leu Arg Ile Ser Leu Ala Ile Val Met Leu Leu Met Thr Val Val Gly
 35 40 45

Phe Leu Gly Asn Thr Val Val Cys Ile Ile Val Tyr Gln Arg Pro Ala
 50 55 60

Met Arg Ser Ala Ile Asn Leu Leu Leu Ala Thr Leu Ala Phe Ser Asp
 65 70 75 80

Ile Met Leu Ser Leu Cys Cys Met Pro Phe Thr Ala Val Thr Leu Ile
 85 90 95

Thr Val Arg Trp His Phe Gly Asp His Phe Cys Arg Leu Ser Ala Thr
 100 105 110

Leu Tyr Trp Phe Phe Val Leu Glu Gly Val Ala Ile Leu Leu Ile Ile
 115 120 125

Ser Val Asp Arg Phe Leu Ile Ile Val Gln Arg Gln Asp Lys Leu Asn
 130 135 140

Pro Arg Arg Ala Lys Val Ile Ile Ala Val Ser Trp Val Leu Ser Phe
 145 150 155 160

Cys Ile Ala Gly Pro Ser Leu Thr Gly Trp Thr Leu Val Glu Val Pro
 165 170 175

Ala Arg Ala Pro Gln Cys Val Leu Gly Tyr Thr Glu Leu Pro Ala Asp
 180 185 190

Arg Ala Tyr Val Val Thr Leu Val Val Ala Val Phe Phe Ala Pro Phe
 195 200 205

Gly Val Met Leu Cys Ala Tyr Met Cys Ile Leu Asn Thr Val Arg Lys

```
<210> 7
<211> 2748
<212> DNA
<213> Homo. sapiens
```

```

<400> 7
atggtccagc tgaggaagct gctccgcgct ctgactttga tgaagtccc ctgctgcgtg 60
ctggaggtgc tcctgtgctc gctggcggcg gcggcgcgcg gccaggagat gtacgccccg 120
cactcaatcc ggatcgaggg ggacgtcacc ctcggggggc tgttccccgt gcacgccaag 180
ggtcccagcg gagtgcctg cggcgacatc aagagggaaa acgggatcca caggctggaa 240

```

gcgatgctct acgccctgga ccagatcaac agtgatccca acctactgcc caacgtgacg	300
ctgggcgcg cgatcctgga cacttggtcc agggacactt acgcgctcga acagtcgctt	360
actttcgtcc aggcgctcat ccagaaggac acctccgacg tgcgctgcac caacggcgaa	420
ccgccggttt tcgtcaagcc ggagaaagta gttggagtga ttggggcttc ggggagttcg	480
gtctccatca tggtagccaa catcctgagg ctcttcaga tccccagat tagttatgca	540
tcaacggcac ccgagctaag tgatgaccgg cgctatgact tcttctctcg cgtggtgcca	600
cccgattcct tccaagccca ggccatggtg gacattgtaa aggccctagg ctggaattat	660
gtgtctaccc tcgcatcgga aggaagtatt ggagagaaag gtgtggagtc cttcacgcag	720
atttccaaag aggcaggtgg actctgcatt gccagtcgg tgagaatccc ccaggaacgc	780
aaagacagga ccattgactt tgatagaatt atcaaacagc tcctggacac cccaactcc	840
agggccgctg tgatttttgc caacgatgag gatataaagc agatccttgc agcagccaaa	900
agagctgacc aagttggcca ttttcttgg gtgggatcag acagctgggg atccaaaata	960
aaccactgc accagcatga agatatcgca gaaggggcca tcaccattca gcccaagcga	1020
gccacggtgg aagggtttga tgcctacttt acgtcccgtg cacttgaaaa caacagaaga	1080
aatgtatggt ttgccgaata ctgggaggaa aacttcaact gcaagttgac gattagtggg	1140
tcaaaaaaag aagacacaga tcgcaaatgc acaggacagg agagaattgg aaaagattcc	1200
aactatgagc agggggtaa agtccagttc gtgattgacg cagtctatgc tatggctcac	1260
gcccttcacc acatgaacaa ggatctctgt gctgactacc ggggtgtctg ccagagatg	1320
gagcaagctg gaggcaagaa gttgctgaag tataacgca atgttaattt caatggtagt	1380
gctggcactc cagtgatgtt taacaagaac ggggatgcac ctgggcgtta tgacatcttt	1440
cagtaccaga ccacaaacac cagcaacccg ggttaccgtc tgatcgggca gtggacagac	1500
gaacttcagc tcaatataga agacatgcag tggggtaaag ggtccgaga gatacccgcc	1560
tcagtgtgca cactaccatg taagccagga cagagaaaga agacacagaa aggaactcct	1620
tgctgttga cctgtgagcc ttgcgatggt taccagtacc agtttgatga gatgacatgc	1680
cagcattgcc cctatgacca gagggccaat gaaaatcgaa ccggatgcca ggatattccc	1740
atcatcaaac tggagtggca ctccccctgg gctgtgattc ctgtcttctt ggcaatgttg	1800
gggatcattg ccaccatctt tgtcatggcc actttcatcc gctacaatga cacgcccatt	1860
gtccgggcat ctgggcggga actcagctat gttcttttga cgggcatctt tctttgctac	1920
atcatcactt tcctgatgat tgccaaacca gatgtggcag tgtgttcttt ccggcgagtt	1980

ttcttgggct tgggtatgtg catcagttat gcagccctct tgacgaaaac aaatcggatt 2040
 tategcatat ttgagcaggg caagaaatca gtaacagctc ccagactcat aagcccaaca 2100
 tcacaactgg caatcacttc cagtttaata tcagttcagc ttctaggggt gttcatttgg 2160
 tttggtgttg atccacccaa catcatcata gactacgatg aacacaagac aatgaaccct 2220
 gagcaagcca gaggggttct caagtgtgac attacagatc tccaaatcat ttgctccttg 2280
 ggatatagca ttcttctcat ggtcacatgt actgtgtatg ccatcaagac tcgggggtgta 2340
 cccgagaatt ttaacgaagc caagcccatt ggattcacta tgtacacgac atgtatagta 2400
 tggcttgctt tcattccaat tttttttggc accgctcaat cagcggaaaa gctctacata 2460
 caaactacca cgcttacaat ctccatgaac ctaagtgcac cagtggcgct ggggatgcta 2520
 tacatgccga aagtgtacat catcattttc caccctgaac tcaatgtcca gaaacggaag 2580
 cgaagcttca aggcggtagt cacagcagcc accatgtcat cgaggctgtc acacaaaccc 2640
 agtgacagac ccaacgggtga ggcaaagacc gagctctgtg aaaacgtaga cccaaacagc 2700
 cctgctgcaa aaaagaagta tgtcagttat aataacctgg ttatctaa 2748

<210> 8
 <211> 915
 <212> PRT
 <213> Homo sapiens

<400> 8

Met Val Gln Leu Arg Lys Leu Leu Arg Val Leu Thr Leu Met Lys Phe
 1 5 10 15

Pro Cys Cys Val Leu Glu Val Leu Leu Cys Ala Leu Ala Ala Ala
 20 25 30

Arg Gly Gln Glu Met Tyr Ala Pro His Ser Ile Arg Ile Glu Gly Asp
 35 40 45

Val Thr Leu Gly Gly Leu Phe Pro Val His Ala Lys Gly Pro Ser Gly
 50 55 60

Val Pro Cys Gly Asp Ile Lys Arg Glu Asn Gly Ile His Arg Leu Glu
 65 70 75 80

Ala Met Leu Tyr Ala Leu Asp Gln Ile Asn Ser Asp Pro Asn Leu Leu
 85 90 95

Pro Asn Val Thr Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp
 100 105 110

Thr Tyr Ala Leu Glu Gln Ser Leu Thr Phe Val Gln Ala Leu Ile Gln
 115 120 125

Lys Asp Thr Ser Asp Val Arg Cys Thr Asn Gly Glu Pro Pro Val Phe
 130 135 140

Val Lys Pro Glu Lys Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser
 145 150 155 160

Val Ser Ile Met Val Ala Asn Ile Leu Arg Leu Phe Gln Ile Pro Gln
 165 170 175

Ile Ser Tyr Ala Ser Thr Ala Pro Glu Leu Ser Asp Asp Arg Arg Tyr
 180 185 190

Asp Phe Phe Ser Arg Val Val Pro Pro Asp Ser Phe Gln Ala Gln Ala
 195 200 205

Met Val Asp Ile Val Lys Ala Leu Gly Trp Asn Tyr Val Ser Thr Leu
 210 215 220

Ala Ser Glu Gly Ser Tyr Gly Glu Lys Gly Val Glu Ser Phe Thr Gln
 225 230 235 240

Ile Ser Lys Glu Ala Gly Gly Leu Cys Ile Ala Gln Ser Val Arg Ile
 245 250 255

Pro Gln Glu Arg Lys Asp Arg Thr Ile Asp Phe Asp Arg Ile Ile Lys
 260 265 270

Gln Leu Leu Asp Thr Pro Asn Ser Arg Ala Val Val Ile Phe Ala Asn
 275 280 285

Asp Glu Asp Ile Lys Gln Ile Leu Ala Ala Ala Lys Arg Ala Asp Gln
 290 295 300

Val Gly His Phe Leu Trp Val Gly Ser Asp Ser Trp Gly Ser Lys Ile
 305 310 315 320

Asn Pro Leu His Gln His Glu Asp Ile Ala Glu Gly Ala Ile Thr Ile
 325 330 335

Gln Pro Lys Arg Ala Thr Val Glu Gly Phe Asp Ala Tyr Phe Thr Ser
 340 345 350

Arg Thr Leu Glu Asn Asn Arg Arg Asn Val Trp Phe Ala Glu Tyr Trp
 355 360 365

Glu Glu Asn Phe Asn Cys Lys Leu Thr Ile Ser Gly Ser Lys Lys Glu
 370 375 380

Asp Thr Asp Arg Lys Cys Thr Gly Gln Glu Arg Ile Gly Lys Asp Ser
 385 390 395 400

Asn Tyr Glu Gln Glu Gly Lys Val Gln Phe Val Ile Asp Ala Val Tyr
 405 410 415

Ala Met Ala His Ala Leu His His Met Asn Lys Asp Leu Cys Ala Asp
 420 425 430

Tyr Arg Gly Val Cys Pro Glu Met Glu Gln Ala Gly Gly Lys Lys Leu
 435 440 445

Leu Lys Tyr Ile Arg Asn Val Asn Phe Asn Gly Ser Ala Gly Thr Pro
 450 455 460

Val Met Phe Asn Lys Asn Gly Asp Ala Pro Gly Arg Tyr Asp Ile Phe
 465 470 475 480

Gln Tyr Gln Thr Thr Asn Thr Ser Asn Pro Gly Tyr Arg Leu Ile Gly
 485 490 495

Gln Trp Thr Asp Glu Leu Gln Leu Asn Ile Glu Asp Met Gln Trp Gly
 500 505 510

Lys Gly Val Arg Glu Ile Pro Ala Ser Val Cys Thr Leu Pro Cys Lys
 515 520 525

Pro Gly Gln Arg Lys Lys Thr Gln Lys Gly Thr Pro Cys Cys Trp Thr
 530 535 540

Cys Glu Pro Cys Asp Gly Tyr Gln Tyr Gln Phe Asp Glu Met Thr Cys
 545 550 555 560

Gln His Cys Pro Tyr Asp Gln Arg Pro Asn Glu Asn Arg Thr Gly Cys
 565 570 575

Gln Asp Ile Pro Ile Ile Lys Leu Glu Trp His Ser Pro Trp Ala Val
 580 585 590

Ile Pro Val Phe Leu Ala Met Leu Gly Ile Ile Ala Thr Ile Phe Val
 595 600 605

Met Ala Thr Phe Ile Arg Tyr Asn Asp Thr Pro Ile Val Arg Ala Ser
 610 615 620

Gly Arg Glu Leu Ser Tyr Val Leu Leu Thr Gly Ile Phe Leu Cys Tyr
 625 630 635 640

Ile Ile Thr Phe Leu Met Ile Ala Lys Pro Asp Val Ala Val Cys Ser
 645 650 655

Phe Arg Arg Val Phe Leu Gly Leu Gly Met Cys Ile Ser Tyr Ala Ala
 660 665 670

Leu Leu Thr Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys
 675 680 685

Lys Ser Val Thr Ala Pro Arg Leu Ile Ser Pro Thr Ser Gln Leu Ala
 690 695 700

Ile Thr Ser Ser Leu Ile Ser Val Gln Leu Leu Gly Val Phe Ile Trp
 705 710 715 720

Phe Gly Val Asp Pro Pro Asn Ile Ile Ile Asp Tyr Asp Glu His Lys
 725 730 735

Thr Met Asn Pro Glu Gln Ala Arg Gly Val Leu Lys Cys Asp Ile Thr
 740 745 750

Asp Leu Gln Ile Ile Cys Ser Leu Gly Tyr Ser Ile Leu Leu Met Val
 755 760 765

Thr Cys Thr Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Asn Phe

770 775 780
 Asn Glu Ala Lys Pro Ile Gly Phe Thr Met Tyr Thr Thr Cys Ile Val
 785 790 795 800
 Trp Leu Ala Phe Ile Pro Ile Phe Phe Gly Thr Ala Gln Ser Ala Glu
 805 810 815
 Lys Leu Tyr Ile Gln Thr Thr Thr Leu Thr Ile Ser Met Asn Leu Ser
 820 825 830
 Ala Ser Val Ala Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile
 835 840 845
 Ile Phe His Pro Glu Leu Asn Val Gln Lys Arg Lys Arg Ser Phe Lys
 850 855 860
 Ala Val Val Thr Ala Ala Thr Met Ser Ser Arg Leu Ser His Lys Pro
 865 870 875 880
 Ser Asp Arg Pro Asn Gly Glu Ala Lys Thr Glu Leu Cys Glu Asn Val
 885 890 895
 Asp Pro Asn Ser Pro Ala Ala Lys Lys Lys Tyr Val Ser Tyr Asn Asn
 900 905 910
 Leu Val Ile
 915

<210> 9
 <211> 1842
 <212> DNA
 <213> Homo sapiens

<400> 9
 atgcgagccc cgggcgcgct tctcgcccgc atgtcgcggc tactgcttct gctactgctc 60
 aagggtgtctg cctcttctgc cctcggggtc gcccttgcgt ccagaaacga aacttgtctg 120
 ggggagagct gtgcacctac agtgatccag cgccgcggca gggacgcctg gggaccggga 180
 aattctgcaa gagacgttct gcgagcccga gcaccaggga aggagcaggg ggcagcgttt 240
 cttgcccggac cctcctggga cctgccggcg gccccggggc gtgaccgggc tgcaggcaga 300
 ggggcggaag cgtcggcagc cggacccccg ggacctccaa ccaggccacc tggccccttg 360

```

aggtggaaag gtgctcgggg tcaggagcct tctgaaactt tggggagagg gaacccacg 420
gccctccagc tcttccttca gatctcagag gaggaagaga aggggtcccag aggcgctggc 480
atttccgggc gtagccagga gcagagtgtg aagacagtcc ccggagccag cgatcttttt 540
tactggccaa ggagagccgg gaaactccag gggtcccacc acaagcccct gtccaagacg 600
gccaatggac tggcggggca cgaagggtgg acaattgcac tcccgggccg ggcgctggcc 660
cagaatggat ccttgggtga aggaatccat gagcctgggg gtccccgccg gggaaacagc 720
acgaaccggc gtgtgagact gaagaacccc ttctaccgcg tgaccagga gtcctatgga 780
gcctacgcgg tcatgtgtct gtccgtgggt atcttcggga ccggcatcat tggcaacctg 840
gcggtgatgt gcatcgtgtg ccacaactac tacatgcgga gcatctcaa ctccctcttg 900
gccaaactgg ccttctggga ctttctcatc atcttcttct gccttcgct ggatcatctc 960
cacgagctga ccaagaagtg gctgctggag gacttctcct gcaagatcgt gccctatata 1020
gaggtcgctt ctctgggagt caccaccttc accttatgtg ctctgtgcat agaccgcttc 1080
cgtgctgcc acaacgtaca gatgtactac gaaatgatcg aaaactgttc ctcaacaact 1140
gccaaacttg ctgttatatg ggtgggagct ctattgttag cacttcaga agttgttctc 1200
cgccagctga gcaaggagga tttggggttt agtggccgag ctccggcaga aaggtgcatt 1260
attaagatct ctccctgattt accagacacc atctatgttc tagccctcac ctacgacagt 1320
gcgagactgt ggtggtatatt tggctgttac ttttgtttgc ccacgctttt caccatcacc 1380
tgctctctag tgactgcgag gaaaatccgc aaagcagaga aagcctgtac ccgagggaat 1440
aaacggcaga ttcaactaga gagtcagatg aactgtacag tagtggcact gaccatttta 1500
tatggatttt gcattattcc tgaaaatata tgcaacattg ttactgccta catggctaca 1560
ggggtttcac agcagacaat ggacctcctt aatatcatca gccagttcct tttgttcttt 1620
aagtccgtgt tcaccccagt cctccttttc tgtctctgca aacccttcag tcgggccttc 1680
atggagtgtc gctgctgttg ctgtgaggaa tgcattcaga agtcttcaac ggtgaccagt 1740
gatgacaatg acaacgagta caccacggaa ctcgaactct cgcctttcag taccatacgc 1800
cgtgaaatgt ccacttttgc ttctgtcgga actcattgct ga 1842

```

```

<210> 10
<211> 613
<212> PRT
<213> Homo sapiens
<400> 10

```

Met Arg Ala Pro Gly Ala Leu Leu Ala Arg Met Ser Arg Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Leu Lys Val Ser Ala Ser Ser Ala Leu Gly Val Ala Pro
 20 25 30
 Ala Ser Arg Asn Glu Thr Cys Leu Gly Glu Ser Cys Ala Pro Thr Val
 35 40 45
 Ile Gln Arg Arg Gly Arg Asp Ala Trp Gly Pro Gly Asn Ser Ala Arg
 50 55 60
 Asp Val Leu Arg Ala Arg Ala Pro Arg Glu Glu Gln Gly Ala Ala Phe
 65 70 75 80
 Leu Ala Gly Pro Ser Trp Asp Leu Pro Ala Ala Pro Gly Arg Asp Pro
 85 90 95
 Ala Ala Gly Arg Gly Ala Glu Ala Ser Ala Ala Gly Pro Pro Gly Pro
 100 105 110
 Pro Thr Arg Pro Pro Gly Pro Trp Arg Trp Lys Gly Ala Arg Gly Gln
 115 120 125
 Glu Pro Ser Glu Thr Leu Gly Arg Gly Asn Pro Thr Ala Leu Gln Leu
 130 135 140
 Phe Leu Gln Ile Ser Glu Glu Glu Glu Lys Gly Pro Arg Gly Ala Gly
 145 150 155 160
 Ile Ser Gly Arg Ser Gln Glu Gln Ser Val Lys Thr Val Pro Gly Ala
 165 170 175
 Ser Asp Leu Phe Tyr Trp Pro Arg Arg Ala Gly Lys Leu Gln Gly Ser
 180 185 190
 His His Lys Pro Leu Ser Lys Thr Ala Asn Gly Leu Ala Gly His Glu
 195 200 205
 Gly Trp Thr Ile Ala Leu Pro Gly Arg Ala Leu Ala Gln Asn Gly Ser
 210 215 220

Leu Gly Glu Gly Ile His Glu Pro Gly Gly Pro Arg Arg Gly Asn Ser
 225 230 235 240

Thr Asn Arg Arg Val Arg Leu Lys Asn Pro Phe Tyr Pro Leu Thr Gln
 245 250 255

Glu Ser Tyr Gly Ala Tyr Ala Val Met Cys Leu Ser Val Val Ile Phe
 260 265 270

Gly Thr Gly Ile Ile Gly Asn Leu Ala Val Met Cys Ile Val Cys His
 275 280 285

Asn Tyr Tyr Met Arg Ser Ile Ser Asn Ser Leu Leu Ala Asn Leu Ala
 290 295 300

Phe Trp Asp Phe Leu Ile Ile Phe Phe Cys Leu Pro Leu Val Ile Phe
 305 310 315 320

His Glu Leu Thr Lys Lys Trp Leu Leu Glu Asp Phe Ser Cys Lys Ile
 325 330 335

Val Pro Tyr Ile Glu Val Ala Ser Leu Gly Val Thr Thr Phe Thr Leu
 340 345 350

Cys Ala Leu Cys Ile Asp Arg Phe Arg Ala Ala Thr Asn Val Gln Met
 355 360 365

Tyr Tyr Glu Met Ile Glu Asn Cys Ser Ser Thr Thr Ala Lys Leu Ala
 370 375 380

Val Ile Trp Val Gly Ala Leu Leu Leu Ala Leu Pro Glu Val Val Leu
 385 390 395 400

Arg Gln Leu Ser Lys Glu Asp Leu Gly Phe Ser Gly Arg Ala Pro Ala
 405 410 415

Glu Arg Cys Ile Ile Lys Ile Ser Pro Asp Leu Pro Asp Thr Ile Tyr
 420 425 430

Val Leu Ala Leu Thr Tyr Asp Ser Ala Arg Leu Trp Trp Tyr Phe Gly
 435 440 445

Cys Tyr Phe Cys Leu Pro Thr Leu Phe Thr Ile Thr Cys Ser Leu Val

450 455 460
 Thr Ala Arg Lys Ile Arg Lys Ala Glu Lys Ala Cys Thr Arg Gly Asn
 465 470 475 480
 Lys Arg Gln Ile Gln Leu Glu Ser Gln Met Asn Cys Thr Val Val Ala
 485 490 495
 Leu Thr Ile Leu Tyr Gly Phe Cys Ile Ile Pro Glu Asn Ile Cys Asn
 500 505 510
 Ile Val Thr Ala Tyr Met Ala Thr Gly Val Ser Gln Gln Thr Met Asp
 515 520 525
 Leu Leu Asn Ile Ile Ser Gln Phe Leu Leu Phe Phe Lys Ser Cys Val
 530 535 540
 Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe
 545 550 555 560
 Met Glu Cys Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser
 565 570 575
 Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu
 580 585 590
 Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser
 595 600 605
 Val Gly Thr His Cys
 610

<210> 11
 <211> 1086
 <212> DNA
 <213> Homo sapiens

<400> 11
 atgtcccctg aatgcgcgcg ggcagcgggc gacgcgccct tgcgcagcct ggagcaagcc 60
 aaccgcaccc gctttccctt cttctccgac gtcaagggcg accaccggct ggtgctggcc 120
 gcgggtggaga caaccgtgct ggtgctcatc ttgcagtgt cgctgctggg caacgtgtgc 180
 gccctggtgc tggtggcgcg ccgacgacgc cgcggcgcga ctgcctgcct ggtactcaac 240

```

ctcttctgcg cggacctgct cttcatcagc gctatccctc tgggtgctggc cgtgcgctgg 300
actgaggcct ggctgctggg ccccgttgcc tgccacctgc tcttctacgt gatgacctg 360
agcggcagcg tcaccatcct cacgctggcc gcggtcagcc tggagcgcat ggtgtgcatc 420
gtgcacctgc agcgcggcgt gcggggctct gggcggcggg cgcgggcagt gctgctggcg 480
ctcatctggg gctattcggc ggtcgccgct ctgcctctct gcgtcttctt tcgagtcgtc 540
ccgcaacggc tccccggcgc cgaccaggaa atttcgattt gcacactgat ttggcccacc 600
attcctggag agatctcgtg ggatgtctct tttgttactt tgaacttctt ggtgccagga 660
ctgggtcattg tgatcagtta ctccaaaatt ttacagatca caaaggcatc aaggaagagg 720
ctcacggtaa gcctggccta ctcgagagac caccagatcc gcgtgtccca gcaggacttc 780
cggtctcttc gcacctctt cctcctcatg gtctccttct tcatcatgtg gagccccatc 840
atcatcacca tctcctcat cctgatccag aacttcaagc aagacctggt catctggccg 900
tccctcttct tctgggtggt ggccttcaca tttgctaatt cagccctaaa ccccatcctc 960
tacaacatga cactgtgcag gaatgagtgg aagaaaattt tttgctgctt ctggttccca 1020
gaaaagggag ccattttaac agacacatct gtcaaaagaa atgacttgtc gattatttct 1080
ggctaa 1086

```

<210> 12
 <211> 361
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 12

Met Ser Pro Glu Cys Ala Arg Ala Ala Gly Asp Ala Pro Leu Arg Ser
 1 5 10 15

Leu Glu Gln Ala Asn Arg Thr Arg Phe Pro Phe Phe Ser Asp Val Lys
 20 25 30

Gly Asp His Arg Leu Val Leu Ala Ala Val Glu Thr Thr Val Leu Val
 35 40 45

Leu Ile Phe Ala Val Ser Leu Leu Gly Asn Val Cys Ala Leu Val Leu
 50 55 60

Val Ala Arg Arg Arg Arg Arg Gly Ala Thr Ala Cys Leu Val Leu Asn
 65 70 75 80

Leu Phe Cys Ala Asp Leu Leu Phe Ile Ser Ala Ile Pro Leu Val Leu
 85 90 95

Ala Val Arg Trp Thr Glu Ala Trp Leu Leu Gly Pro Val Ala Cys His
 100 105 110

Leu Leu Phe Tyr Val Met Thr Leu Ser Gly Ser Val Thr Ile Leu Thr
 115 120 125

Leu Ala Ala Val Ser Leu Glu Arg Met Val Cys Ile Val His Leu Gln
 130 135 140

Arg Gly Val Arg Gly Pro Gly Arg Arg Ala Arg Ala Val Leu Leu Ala
 145 150 155 160

Leu Ile Trp Gly Tyr Ser Ala Val Ala Ala Leu Pro Leu Cys Val Phe
 165 170 175

Phe Arg Val Val Pro Gln Arg Leu Pro Gly Ala Asp Gln Glu Ile Ser
 180 185 190

Ile Cys Thr Leu Ile Trp Pro Thr Ile Pro Gly Glu Ile Ser Trp Asp
 195 200 205

Val Ser Phe Val Thr Leu Asn Phe Leu Val Pro Gly Leu Val Ile Val
 210 215 220

Ile Ser Tyr Ser Lys Ile Leu Gln Ile Thr Lys Ala Ser Arg Lys Arg
 225 230 235 240

Leu Thr Val Ser Leu Ala Tyr Ser Glu Ser His Gln Ile Arg Val Ser
 245 250 255

Gln Gln Asp Phe Arg Leu Phe Arg Thr Leu Phe Leu Leu Met Val Ser
 260 265 270

Phe Phe Ile Met Trp Ser Pro Ile Ile Ile Thr Ile Leu Leu Ile Leu
 275 280 285

Ile Gln Asn Phe Lys Gln Asp Leu Val Ile Trp Pro Ser Leu Phe Phe

290

295

300

Trp Val Val Ala Phe Thr Phe Ala Asn Ser Ala Leu Asn Pro Ile Leu
 305 310 315 320

Tyr Asn Met Thr Leu Cys Arg Asn Glu Trp Lys Lys Ile Phe Cys Cys
 325 330 335

Phe Trp Phe Pro Glu Lys Gly Ala Ile Leu Thr Asp Thr Ser Val Lys
 340 345 350

Arg Asn Asp Leu Ser Ile Ile Ser Gly
 355 360

<210> 13
 <211> 1212
 <212> DNA
 <213> Homo sapiens

<400> 13
 atggccttgca atggcagtgcc ggccaggggg cactttgacc ctgaggactt gaacctgact 60
 gacgaggcac tgagactcaa gtacctgggg cccagcaga cagagctgtt catgcccac 120
 tgtgccacat acctgctgat cttcgtggtg ggcgctgtgg gcaatgggct gacctgtctg 180
 gtcacctcgc gccacaaggc catgcgcacg cctaccaact actacctctt cagcctggcc 240
 gtgtcggacc tgctggtgct gctgggtggc ctgcccctgg agctctatga gatgtggcac 300
 aactaccctt tctgctggg cgttggtggc tgctatttcc gcacgctact gtttgagatg 360
 gtctgcctgg cctcagtgtt caacgtcact gccctgagcg tggaaacgcta tgtggccgtg 420
 gtgcacccac tccaggccag gtccatggtg acgcggggcc atgtgcgccg agtgcttggg 480
 gccgtctggg gtcttgccat gctctgtctc ctgccaaca ccagcctgca cggcatccgg 540
 cagctgcacg tgccctgccg gggcccagtg ccagactcag ctgtttgcat gctgggtccg 600
 ccacggggcc tctacaacat ggtagtgcag accaccggcg tgctcttctt ctgcctgccc 660
 atggccatca tgagcgtgct ctacctgtc attgggtgc gactgcggcg ggagaggctg 720
 ctgctcatgc aggaggccaa gggcaggggc tctgcagcag ccaggctccag atacacctgc 780
 aggctccagc agcacgatcg gggccggaga caagtgacca agatgctgtt tgcctggtc 840
 gtggtgtttg gcatctgtg gggcccgttc cacgccgacc gcgtcatgtg gagcgtcgtg 900
 tcacagtga cagatggcct gcacctggcc ttccagcacg tgcacgtcat ctccggcatc 960

```

ttcttctacc tgggctcggc ggccaacccc gtgctctata gcctcatgtc cagccgcttc 1020
cgagagacct tccaggaggc cctgtgcctc ggggcctgct gccatcgctt cagaccccgcc 1080
cacagctccc acagcctcag caggatgacc acaggcagca ccctgtgtga tgtgggctcc 1140
ctgggcagct gggtcacccc cctggctggg aacgatggcc cagaggcgca gcaagagacc 1200
gatccatcct ga 1212

```

```

<210> 14
<211> 403
<212> PRT
<213> Homo sapiens

```

```

<400> 14

```

```

Met Ala Cys Asn Gly Ser Ala Ala Arg Gly His Phe Asp Pro Glu Asp
1          5          10          15

```

```

Leu Asn Leu Thr Asp Glu Ala Leu Arg Leu Lys Tyr Leu Gly Pro Gln
          20          25          30

```

```

Gln Thr Glu Leu Phe Met Pro Ile Cys Ala Thr Tyr Leu Leu Ile Phe
          35          40          45

```

```

Val Val Gly Ala Val Gly Asn Gly Leu Thr Cys Leu Val Ile Leu Arg
50          55          60

```

```

His Lys Ala Met Arg Thr Pro Thr Asn Tyr Tyr Leu Phe Ser Leu Ala
65          70          75          80

```

```

Val Ser Asp Leu Leu Val Leu Leu Val Gly Leu Pro Leu Glu Leu Tyr
          85          90          95

```

```

Glu Met Trp His Asn Tyr Pro Phe Leu Leu Gly Val Gly Gly Cys Tyr
          100          105          110

```

```

Phe Arg Thr Leu Leu Phe Glu Met Val Cys Leu Ala Ser Val Leu Asn
          115          120          125

```

```

Val Thr Ala Leu Ser Val Glu Arg Tyr Val Ala Val Val His Pro Leu
          130          135          140

```

```

Gln Ala Arg Ser Met Val Thr Arg Ala His Val Arg Arg Val Leu Gly
145          150          155          160

```

Ala Val Trp Gly Leu Ala Met Leu Cys Ser Leu Pro Asn Thr Ser Leu
 165 170 175

His Gly Ile Arg Gln Leu His Val Pro Cys Arg Gly Pro Val Pro Asp
 180 185 190

Ser Ala Val Cys Met Leu Val Arg Pro Arg Ala Leu Tyr Asn Met Val
 195 200 205

Val Gln Thr Thr Ala Leu Leu Phe Phe Cys Leu Pro Met Ala Ile Met
 210 215 220

Ser Val Leu Tyr Leu Leu Ile Gly Leu Arg Leu Arg Arg Glu Arg Leu
 225 230 235 240

Leu Leu Met Gln Glu Ala Lys Gly Arg Gly Ser Ala Ala Ala Arg Ser
 245 250 255

Arg Tyr Thr Cys Arg Leu Gln Gln His Asp Arg Gly Arg Arg Gln Val
 260 265 270

Thr Lys Met Leu Phe Val Leu Val Val Val Phe Gly Ile Cys Trp Ala
 275 280 285

Pro Phe His Ala Asp Arg Val Met Trp Ser Val Val Ser Gln Trp Thr
 290 295 300

Asp Gly Leu His Leu Ala Phe Gln His Val His Val Ile Ser Gly Ile
 305 310 315 320

Phe Phe Tyr Leu Gly Ser Ala Ala Asn Pro Val Leu Tyr Ser Leu Met
 325 330 335

Ser Ser Arg Phe Arg Glu Thr Phe Gln Glu Ala Leu Cys Leu Gly Ala
 340 345 350

Cys Cys His Arg Leu Arg Pro Arg His Ser Ser His Ser Leu Ser Arg
 355 360 365

Met Thr Thr Gly Ser Thr Leu Cys Asp Val Gly Ser Leu Gly Ser Trp
 370 375 380

Val His Pro Leu Ala Gly Asn Asp Gly Pro Glu Ala Gln Gln Glu Thr
 385 390 395 400

Asp Pro Ser

<210> 15
 <211> 930
 <212> DNA
 <213> Homo sapiens

<400> 15
 atgaatggca cctacaacac ctgtggctcc agcgacctca cctggccccc agcgatcaag 60
 ctgggcttct acgcctactt gggcgctcctg ctggtgctag gcctgctgct caacagcctg 120
 gcgctctggg tgttctgctg ccgcatgcag cagtggacgg agaccgcgat ctacatgacc 180
 aacctggcgg tggccgacct ctgcctgctg tgcaccttgc ccttcgtgct gcaactccctg 240
 cgagacacct cagacacgcc gctgtgccag ctctcccagg gcatctacct gaccaacagg 300
 tacatgagca tcagcctggg caccggccatc gccgtggacc gctatgtggc cgtgcggcac 360
 ccgctgcgtg cccgcgggct gcggtccccc aggcaggctg cggccggtgt gcggttcctc 420
 tgggtgctgg tcatcggtc cctgggtggc cgctggctcc tggggattca ggaggggcggc 480
 ttctgcttca ggagcaccgg gcacaatttc aactccatgc ggttcccgt gctgggattc 540
 tacctgcccc tggccgtggg ggtcttctgc tccctgaagg tggtgactgc cctggcccag 600
 aggccaccca ccgacgtggg gcaggcagag gccacccgca aggctgcccg catggtctgg 660
 gccaacctcc tgggtgtcgt ggtctgcttc ctgcccctgc acgtggggct gacagtgcgc 720
 ctgcgagtgg gctggaacgc ctgtgccctc ctggagacga tccgtcgcgc cctgtacata 780
 accagcaagc tctcagatgc caactgctgc ctggacgcca tctgctacta ctacatggcc 840
 aaggagtcc aggaggcgtc tgcaactggc gtggctcccc gtgctaaggc ccacaaaagc 900
 caggactctc tgtcgtgac cctcgcctaa 930

<210> 16
 <211> 309
 <212> PRT
 <213> Homo sapiens

<400> 16

Met Asn Gly Thr Tyr Asn Thr Cys Gly Ser Ser Asp Leu Thr Trp Pro
 1 5 10 15

Pro Ala Ile Lys Leu Gly Phe Tyr Ala Tyr Leu Gly Val Leu Leu Val
 20 25 30

Leu Gly Leu Leu Leu Asn Ser Leu Ala Leu Trp Val Phe Cys Cys Arg
 35 40 45

Met Gln Gln Trp Thr Glu Thr Arg Ile Tyr Met Thr Asn Leu Ala Val
 50 55 60

Ala Asp Leu Cys Leu Leu Cys Thr Leu Pro Phe Val Leu His Ser Leu
 65 70 75 80

Arg Asp Thr Ser Asp Thr Pro Leu Cys Gln Leu Ser Gln Gly Ile Tyr
 85 90 95

Leu Thr Asn Arg Tyr Met Ser Ile Ser Leu Val Thr Ala Ile Ala Val
 100 105 110

Asp Arg Tyr Val Ala Val Arg His Pro Leu Arg Ala Arg Gly Leu Arg
 115 120 125

Ser Pro Arg Gln Ala Ala Ala Val Cys Ala Val Leu Trp Val Leu Val
 130 135 140

Ile Gly Ser Leu Val Ala Arg Trp Leu Leu Gly Ile Gln Glu Gly Gly
 145 150 155 160

Phe Cys Phe Arg Ser Thr Arg His Asn Phe Asn Ser Met Arg Phe Pro
 165 170 175

Leu Leu Gly Phe Tyr Leu Pro Leu Ala Val Val Val Phe Cys Ser Leu
 180 185 190

Lys Val Val Thr Ala Leu Ala Gln Arg Pro Pro Thr Asp Val Gly Gln
 195 200 205

Ala Glu Ala Thr Arg Lys Ala Ala Arg Met Val Trp Ala Asn Leu Leu
 210 215 220

Val Phe Val Val Cys Phe Leu Pro Leu His Val Gly Leu Thr Val Arg
 225 230 235 240

Leu Ala Val Gly Trp Asn Ala Cys Ala Leu Leu Glu Thr Ile Arg Arg
 245 250 255

Ala Leu Tyr Ile Thr Ser Lys Leu Ser Asp Ala Asn Cys Cys Leu Asp
 260 265 270

Ala Ile Cys Tyr Tyr Tyr Met Ala Lys Glu Phe Gln Glu Ala Ser Ala
 275 280 285

Leu Ala Val Ala Pro Arg Ala Lys Ala His Lys Ser Gln Asp Ser Leu
 290 295 300

Cys Val Thr Leu Ala
 305

<210> 17
 <211> 1446
 <212> DNA
 <213> Homo sapiens

<400> 17
 atgCGgtggc tgtggccctt ggctgtctct cttgctgtga ttttggctgt ggggctaagc 60
 agggctctctg ggggtgcccc cctgcacctg ggcaggcaca gagccgagac ccaggagcag 120
 cagagccgat ccaagagggg caccgaggat gaggaggcca agggcgtgca gcagtatgtg 180
 cctgaggagt gggcggagta ccccgggccc attcaccctg ctggcctgca gccaaaccaag 240
 cccttggtgg ccaccagccc taaccccgac aaggatgggg gcaccccgaga cagtgggcag 300
 gaactgaggg gcaatctgac aggggcacca gggcagaggc tacagatcca gaacccctg 360
 tatccggtga ccgagagctc ctacagtgcc tatgccatca tgcttctggc gctggtggtg 420
 tttgcggtgg gcattgtggg caacctgtcg gtcattgtgca tcgtgtggca cagctactac 480
 ctgaagagcg cctggaactc catccttgcc agcctggccc tctgggattt tctggtcctc 540
 tttttctgcc tccctattgt catcttcaac gagatcacca agcagaggct actgggtgac 600
 gtttcttgtc gtgccgtgcc cttcatggag gtctctctc tgggagtcac gactttcagc 660
 ctctgtgccc tgggcattga ccgcttcac gtggccacca gcaccctgcc caaggtaggg 720
 cccatcgagc ggtgccaatc catcctggcc aagttggctg tcatctgggt gggctccatg 780
 acgctggctg tgcctgagct cctgctgtgg cagctggcac aggagcctgc cccaccatg 840
 ggcaccctgg actcatgcat catgaaacct tcagccagcc tgcccgagtc cctgtattca 900
 ctggtgatga cctaccagaa cggccgcatg tgggtgtact ttggctgcta cttctgcctg 960

```

cccatcctct tcacagtcac ctgccagctg gtgacatggc ggggtgcgagg ccctccaggg 1020
aggaagtcag agtgcagggc cagcaagcac gagcagtgtg agagccagct caacagcacc 1080
gtgggtgggcc tgaccgtggt ctacgccttc tgcaccctcc cagagaacgt ctgcaacatc 1140
gtgggtggcct acctctccac cgagctgacc cgccagaccc tggacctcct gggcctcatc 1200
aaccagttct ccaccttctt caagggcgcc atcacccag tgctgctcct ttgcatctgc 1260
aggcgcgtgg gccaggcctt cctggactgc tgctgctgct gctgctgtga ggagtgcggc 1320
ggggcttcgg aggcctctgc tgccaatggg tcggacaaca agctcaagac cgaggtgtcc 1380
tcttccatct acttccacaa gcccaggag tcacccccac tctgcccct gggcacacct 1440
tgctga 1446

```

```

<210> 18
<211> 481
<212> PRT
<213> Homo sapiens

```

```

<400> 18

```

```

Met Arg Trp Leu Trp Pro Leu Ala Val Ser Leu Ala Val Ile Leu Ala
1           5           10           15

```

```

Val Gly Leu Ser Arg Val Ser Gly Gly Ala Pro Leu His Leu Gly Arg
          20           25           30

```

```

His Arg Ala Glu Thr Gln Glu Gln Gln Ser Arg Ser Lys Arg Gly Thr
          35           40           45

```

```

Glu Asp Glu Glu Ala Lys Gly Val Gln Gln Tyr Val Pro Glu Glu Trp
          50           55           60

```

```

Ala Glu Tyr Pro Arg Pro Ile His Pro Ala Gly Leu Gln Pro Thr Lys
65           70           75           80

```

```

Pro Leu Val Ala Thr Ser Pro Asn Pro Asp Lys Asp Gly Gly Thr Pro
          85           90           95

```

```

Asp Ser Gly Gln Glu Leu Arg Gly Asn Leu Thr Gly Ala Pro Gly Gln
          100          105          110

```

```

Arg Leu Gln Ile Gln Asn Pro Leu Tyr Pro Val Thr Glu Ser Ser Tyr
          115          120          125

```

Ser Ala Tyr Ala Ile Met Leu Leu Ala Leu Val Val Phe Ala Val Gly
 130 135 140

Ile Val Gly Asn Leu Ser Val Met Cys Ile Val Trp His Ser Tyr Tyr
 145 150 155 160

Leu Lys Ser Ala Trp Asn Ser Ile Leu Ala Ser Leu Ala Leu Trp Asp
 165 170 175

Phe Leu Val Leu Phe Phe Cys Leu Pro Ile Val Ile Phe Asn Glu Ile
 180 185 190

Thr Lys Gln Arg Leu Leu Gly Asp Val Ser Cys Arg Ala Val Pro Phe
 195 200 205

Met Glu Val Ser Ser Leu Gly Val Thr Thr Phe Ser Leu Cys Ala Leu
 210 215 220

Gly Ile Asp Arg Phe His Val Ala Thr Ser Thr Leu Pro Lys Val Arg
 225 230 235 240

Pro Ile Glu Arg Cys Gln Ser Ile Leu Ala Lys Leu Ala Val Ile Trp
 245 250 255

Val Gly Ser Met Thr Leu Ala Val Pro Glu Leu Leu Leu Trp Gln Leu
 260 265 270

Ala Gln Glu Pro Ala Pro Thr Met Gly Thr Leu Asp Ser Cys Ile Met
 275 280 285

Lys Pro Ser Ala Ser Leu Pro Glu Ser Leu Tyr Ser Leu Val Met Thr
 290 295 300

Tyr Gln Asn Ala Arg Met Trp Trp Tyr Phe Gly Cys Tyr Phe Cys Leu
 305 310 315 320

Pro Ile Leu Phe Thr Val Thr Cys Gln Leu Val Thr Trp Arg Val Arg
 325 330 335

Gly Pro Pro Gly Arg Lys Ser Glu Cys Arg Ala Ser Lys His Glu Gln
 340 345 350

Cys Glu Ser Gln Leu Asn Ser Thr Val Val Gly Leu Thr Val Val Tyr
 355 360 365

Ala Phe Cys Thr Leu Pro Glu Asn Val Cys Asn Ile Val Val Ala Tyr
 370 375 380

Leu Ser Thr Glu Leu Thr Arg Gln Thr Leu Asp Leu Leu Gly Leu Ile
 385 390 395 400

Asn Gln Phe Ser Thr Phe Phe Lys Gly Ala Ile Thr Pro Val Leu Leu
 405 410 415

Leu Cys Ile Cys Arg Pro Leu Gly Gln Ala Phe Leu Asp Cys Cys Cys
 420 425 430

Cys Cys Cys Cys Glu Glu Cys Gly Gly Ala Ser Glu Ala Ser Ala Ala
 435 440 445

Asn Gly Ser Asp Asn Lys Leu Lys Thr Glu Val Ser Ser Ser Ile Tyr
 450 455 460

Phe His Lys Pro Arg Glu Ser Pro Pro Leu Leu Pro Leu Gly Thr Pro
 465 470 475 480

Cys

<210> 19
 <211> 29
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 19
 aaagattcag gtgtgggaag atggaaacc

29

<210> 20
 <211> 29
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 20
aaaggatccc cgacctcaca ttgcttgta

29

<210> 21
<211> 30
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 21
caggaattca tcagaacaga caccatggca

30

<210> 22
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 22
gcaggatcca gagcagtttt ttcgaaacct

31

<210> 23
<211> 33
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 23
tccaagcttc aagggtctct ccacgatggc ctg

33

<210> 24
<211> 33
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 24
tgccaattct ctgtggcccc ctgaccccct aaa

33

<210> 25
<211> 36
<212> DNA
<213> Unknown

<220>

<223> Novel Sequence

<400> 25

ggtaagctta ccatggcctg caacagcacg tccctt

36

<210> 26

<211> 33

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 26

gacgaattca accgcagact ggttttcatt gca

33

<210> 27

<211> 33

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 27

gcaagcttgt gccctcacca agccatgcga gcc

33

<210> 28

<211> 30

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 28

cggaattcag caatgagttc cgacagaagc

30

<210> 29

<211> 37

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 29

accatggctt gcaatggcag tgcggccagg gggcact

37

<210> 30

<211> 39

<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 30
cgaccaggac aaacagcatc ttggtcactt gtctccggc 39

<210> 31
<211> 39
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 31
gaccaagatg ctgtttgtcc tggtcgtggt gtttggcat 39

<210> 32
<211> 35
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 32
cggaattcag gatggatcgg tctcttgctg cgcct 35

<210> 33
<211> 30
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 33
gcgaattccg gctccctgtg ctgccccagg 30

<210> 34
<211> 30
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 34
gcggatcccg gagccccga gacctggccc 30

<210> 35
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 35
ctggaattct cctgctcatc cagccatgcg g 31

<210> 36
<211> 30
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 36
cctggatccc caccctact ggggcctcag 30

<210> 37
<211> 29
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 37
tccagccgtc ccaaactgt cttegctgc 29

<210> 38
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 38
ctccttcggt cctcctatcg ttgtcagaag t 31

<210> 39
<211> 33
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 39
cagaagcaca gatcaaaaaa gatcatcttc ctg 33

<210> 40
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 40
acaggaatca cagccgaggg ggagtgccac t 31

<210> 41
<211> 32
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 41
tgtgttcttt ccggcatgtt ttcttgggt tg 32

<210> 42
<211> 32
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 42
caagcccaag aaaacatgcc ggaaagaaca ca 32

<210> 43
<211> 33
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 43
ctcatggtca catgttgtgt gtatgccatc aag 33

<210> 44
<211> 33
<212> DNA
<213> Unknown

<220>

<223> Novel Sequence

<400> 44

cttgatggca tacacacaac atgtgaccat gag

33

<210> 45

<211> 34

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 45

acgaagccaa gcccaaggga ttcactatgt acac

34

<210> 46

<211> 34

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 46

gtgtacatag tgaatccctt gggcttggt ccgt

34

<210> 47

<211> 35

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 47

gtcaccacct ttcacccgat gtgctctgtg catag

35

<210> 48

<211> 35

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 48

ctatgcacag agcacatcgg gtgaaagggtg gtgac

35

<210> 49

<211> 36

<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 49
ccttttgttc tttaagtcct atgtcacccc agtcct 36

<210> 50
<211> 36
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 50
aggactgggg tgacatagga cttaaagaac aaaagg 36

<210> 51
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 51
atgtggagcc ccatcttcat caccatcctc c 31

<210> 52
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 52
ggaggatggt gatgaagatg gggctccaca t 31

<210> 53
<211> 33
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 53
gccgcggtca gcctgaatcg catggtgtgc atc 33

<210> 54
<211> 33
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 54
gatgcacacc atgcgattca ggctgaccgc ggc 33

<210> 55
<211> 29
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 55
ggccggagac aagtgaaaag atgctgttt 29

<210> 56
<211> 30
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 56
aaacagcatc tttttcactt gtctcggcc 30

<210> 57
<211> 27
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 57
gagagccagc tcaagagcac cgtggtg 27

<210> 58
<211> 31
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 58
 ctccttcggt cctcctatcg ttgtcagaag t 31

<210> 59
 <211> 31
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 59
 agtggcactc cccctcggct gtgattcctg t 31

<210> 60
 <211> 30
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 60
 gccaccgcga aggctaaacg catggtctgg 30

<210> 61
 <211> 31
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 61
 ctccttcggt cctcctatcg ttgtcagaag t 31

<210> 62
 <211> 1062
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 62
 atggaaacca acttctccat tcctctgaat gaaactgagg aggtgctccc tgagcctgct 60
 ggccacaccg ttctgtggat cttctcattg ctagtccacg gagtcacctt tgtcttcggg 120
 gtcctgggca atgggcttgt gatctgggtg gctggattcc ggatgacacg cacagtcaac 180
 accatctgtt acctgaacct ggccctagct gácttctctt tcagtgccat cctaccattc 240

```

cgaatggtct cagtcgccat gagagaaaaa tggccttttg gctcattcct atgtaagtta 300
gttcatgtta tgatagacat caacctgttt gtcagtgtct acctgatcac catcattgct 360
ctggaccgct gtatttgtgt cctgcatcca gcctgggccc agaaccatcg caccatgagt 420
ctggccaaga gggatgatgac gggactctgg attttcacca tagtccttac cttaccaaata 480
ttcatcttct ggactacaat aagtactacg aatggggaca catactgtat tttcaacttt 540
gcattctggg gtgacactgc tgtagagagg ttgaacgtgt tcattaccat ggccaaggtc 600
tttctgatcc tccacttcat tattggcttc agcgtgccta tgtccatcat cacagtctgc 660
tatgggatca tcgctgcaa aattcacaga aaccacatga ttaaattccag ccgtcccaaa 720
cgtgtcttcg ctgctgtggg ggcttctttc ttcactctgtt ggttccctta tgaactaatt 780
ggcattctaa tggcagtctg gctcaaagag atgttggttaa atggcaaata caaaatcatt 840
cttgtctga ttaacccaac aagctccttg gcctttttta acagctgcct caaccaatt 900
ctctacgtct ttatgggtcg taacttccaa gaaagactga ttcgctcttt gccactagt 960
ttggagaggg ccctgactga ggtccctgac tcagcccaga ccagcaacac agacaccact 1020
tctgcttcac ctctgagga gacggagtta caagcaatgt ga 1062

```

<210> 63
 <211> 353
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 63

```

Met Glu Thr Asn Phe Ser Ile Pro Leu Asn Glu Thr Glu Glu Val Leu
1           5           10           15

```

```

Pro Glu Pro Ala Gly His Thr Val Leu Trp Ile Phe Ser Leu Leu Val
          20           25           30

```

```

His Gly Val Thr Phe Val Phe Gly Val Leu Gly Asn Gly Leu Val Ile
          35           40           45

```

```

Trp Val Ala Gly Phe Arg Met Thr Arg Thr Val Asn Thr Ile Cys Tyr
          50           55           60

```

```

Leu Asn Leu Ala Leu Ala Asp Phe Ser Phe Ser Ala Ile Leu Pro Phe
65           70           75           80

```

Arg Met Val Ser Val Ala Met Arg Glu Lys Trp Pro Phe Gly Ser Phe
 85 90 95

Leu Cys Lys Leu Val His Val Met Ile Asp Ile Asn Leu Phe Val Ser
 100 105 110

Val Tyr Leu Ile Thr Ile Ile Ala Leu Asp Arg Cys Ile Cys Val Leu
 115 120 125

His Pro Ala Trp Ala Gln Asn His Arg Thr Met Ser Leu Ala Lys Arg
 130 135 140

Val Met Thr Gly Leu Trp Ile Phe Thr Ile Val Leu Thr Leu Pro Asn
 145 150 155 160

Phe Ile Phe Trp Thr Thr Ile Ser Thr Thr Asn Gly Asp Thr Tyr Cys
 165 170 175

Ile Phe Asn Phe Ala Phe Trp Gly Asp Thr Ala Val Glu Arg Leu Asn
 180 185 190

Val Phe Ile Thr Met Ala Lys Val Phe Leu Ile Leu His Phe Ile Ile
 195 200 205

Gly Phe Ser Val Pro Met Ser Ile Ile Thr Val Cys Tyr Gly Ile Ile
 210 215 220

Ala Ala Lys Ile His Arg Asn His Met Ile Lys Ser Ser Arg Pro Lys
 225 230 235 240

Arg Val Phe Ala Ala Val Val Ala Ser Phe Phe Ile Cys Trp Phe Pro
 245 250 255

Tyr Glu Leu Ile Gly Ile Leu Met Ala Val Trp Leu Lys Glu Met Leu
 260 265 270

Leu Asn Gly Lys Tyr Lys Ile Ile Leu Val Leu Ile Asn Pro Thr Ser
 275 280 285

Ser Leu Ala Phe Phe Asn Ser Cys Leu Asn Pro Ile Leu Tyr Val Phe
 290 295 300

Met Gly Arg Asn Phe Gln Glu Arg Leu Ile Arg Ser Leu Pro Thr Ser
 305 310 315 320

Leu Glu Arg Ala Leu Thr Glu Val Pro Asp Ser Ala Gln Thr Ser Asn
 325 330 335

Thr Asp Thr Thr Ser Ala Ser Pro Pro Glu Glu Thr Glu Leu Gln Ala
 340 345 350

Met

<210> 64
 <211> 1029
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 64
 atggcagagc atgattacca tgaagactat gggttcagca gtttcaatga cagcagccag 60
 gaggagcatc aagccttcct gcagttcagc aaggtctttc tggcctgcat gtacctggtg 120
 gtgtttgtct gtggtctggt ggggaactct ctggtgctgg tcatatccat cttctaccat 180
 aagttgcaga gcctgacgga tgtgttcctg gtgaacctac ccctggctga cctggtgttt 240
 gtctgcactc tggccttctg ggccatgca ggcatccatg aatgggtgtt tggccaggtc 300
 atgtgcaaaa gcctactggg catctacact attaacttct acacgtccat gctcatcctc 360
 acctgcatca ctgtggatcg ttctattgta gtggttaagg ccaccaaggc ctacaaccag 420
 caagccaaga ggatgacctg gggcaaggtc accagcttgc tcatctgggt gatatccctg 480
 ctggtttcct tgcccaaat tatctatggc aatgtottta atctcgacaa gctcatatgt 540
 ggttaccatg acgaggcaat ttccactgtg gttcttgcca cccagatgac actgggggtc 600
 ttcttgccac tgctcaccat gattgtctgc tattcagtca taatcaaaac actgcttcat 660
 gctggaggct tccagaagca cagatcaaaa aagatcatct tcctggtgat ggctgtgttc 720
 ctgctgaccc agatgccctt caacctcatg aagttcatcc gcagcacaca ctgggaatac 780
 tatgccatga ccagctttca ctacaccatc atggtgacag aggccatcgc atacctgagg 840
 gcctgcctta accctgtgct ctatgccttt gtcagcctga agtttcgaaa gaacttctgg 900
 aaacttgtga aggacattgg ttgcctccct taccttgggg totcacatca atggaaatct 960

tctgaggaca attccaagac tttttctgcc tcccacaatg tggaggccac cagcatgttc 1020
cagttatag 1029

<210> 65
<211> 342
<212> PRT
<213> Unknown

<220>
<223> Novel Sequence

<400> 65

Met Ala Glu His Asp Tyr His Glu Asp Tyr Gly Phe Ser Ser Phe Asn
1 5 10 15

Asp Ser Ser Gln Glu Glu His Gln Ala Phe Leu Gln Phe Ser Lys Val
20 25 30

Phe Leu Pro Cys Met Tyr Leu Val Val Phe Val Cys Gly Leu Val Gly
35 40 45

Asn Ser Leu Val Leu Val Ile Ser Ile Phe Tyr His Lys Leu Gln Ser
50 55 60

Leu Thr Asp Val Phe Leu Val Asn Leu Pro Leu Ala Asp Leu Val Phe
65 70 75 80

Val Cys Thr Leu Pro Phe Trp Ala Tyr Ala Gly Ile His Glu Trp Val
85 90 95

Phe Gly Gln Val Met Cys Lys Ser Leu Leu Gly Ile Tyr Thr Ile Asn
100 105 110

Phe Tyr Thr Ser Met Leu Ile Leu Thr Cys Ile Thr Val Asp Arg Phe
115 120 125

Ile Val Val Val Lys Ala Thr Lys Ala Tyr Asn Gln Gln Ala Lys Arg
130 135 140

Met Thr Trp Gly Lys Val Thr Ser Leu Leu Ile Trp Val Ile Ser Leu
145 150 155 160

Leu Val Ser Leu Pro Gln Ile Ile Tyr Gly Asn Val Phe Asn Leu Asp

165 170 175
 Lys Leu Ile Cys Gly Tyr His Asp Glu Ala Ile Ser Thr Val Val Leu
 180 185 190
 Ala Thr Gln Met Thr Leu Gly Phe Phe Leu Pro Leu Leu Thr Met Ile
 195 200 205
 Val Cys Tyr Ser Val Ile Ile Lys Thr Leu Leu His Ala Gly Gly Phe
 210 215 220
 Gln Lys His Arg Ser Lys Lys Ile Ile Phe Leu Val Met Ala Val Phe
 225 230 235 240
 Leu Leu Thr Gln Met Pro Phe Asn Leu Met Lys Phe Ile Arg Ser Thr
 245 250 255
 His Trp Glu Tyr Tyr Ala Met Thr Ser Phe His Tyr Thr Ile Met Val
 260 265 270
 Thr Glu Ala Ile Ala Tyr Leu Arg Ala Cys Leu Asn Pro Val Leu Tyr
 275 280 285
 Ala Phe Val Ser Leu Lys Phe Arg Lys Asn Phe Trp Lys Leu Val Lys
 290 295 300
 Asp Ile Gly Cys Leu Pro Tyr Leu Gly Val Ser His Gln Trp Lys Ser
 305 310 315 320
 Ser Glu Asp Asn Ser Lys Thr Phe Ser Ala Ser His Asn Val Glu Ala
 325 330 335
 Thr Ser Met Phe Gln Leu
 340

<210> 66
 <211> 2748
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 66
 atgggtccagc tgaggaagct gctccgcgtc ctgactttga tgaagttccc ctgctgcgtg

60

ctggaggtgc tcctgtgcgc gctggcggcg gcggcgcgcg gccaggagat gtacgccccg	120
cactcaatcc ggatcgaggg ggacgtcacc ctcggggggc tgttccccgt gcacgccaag	180
ggteccagcg gagtgccctg cggcgacatc aagagggaaa acgggatcca caggctggaa	240
gcgatgctct acgccctgga ccagatcaac agtgatcca acctactgcc caacgtgacg	300
ctgggcgcgc ggatcctgga cacttggttc agggacactt acgcgctcga acagtcgctt	360
actttcgtcc aggcgctcat ccagaaggac acctccgacg tgcgctgcac caacggcgaa	420
ccgccggttt tcgtcaagcc ggagaaagta gttggagtga ttggggcttc ggggagttcg	480
gtctccatca tggtagccaa catcctgagg ctcttcaga tccccagat tagttatgca	540
tcaacggcac ccgagctaag tgatgaccgg cgctatgact tcttctctcg cgtggtgcca	600
cccgattcct tccaagccca ggccatggta gacattgtaa aggcctagg ctggaattat	660
gtgtctaccc tcgcatcgga aggaagttaa ggagagaaag gtgtggagtc cttcacgcag	720
atttccaaag aggcaggtgg actctgcatt gccagtcgg tgagaatccc ccaggaacgc	780
aaagacagga ccattgactt tgatagaatt atcaaacagc tcctggacac cccaactcc	840
agggccgtcg tgatttttgc caacgatgag gatataaagc agatccttgc agcagccaaa	900
agagctgacc aagttggcca ttttctttgg gtgggatcag acagctgggg atccaaaata	960
aaccactgc accagcatga agatatcgca gaaggggcca tcaccattca gcccaagcga	1020
gccacggtgg aagggtttga tgcctacttt acgtcccgta cacttgaaaa caacagaaga	1080
aatgtatggt ttgccgaata ctgggaggaa aacttcaact gcaagttgac gattagtggg	1140
tcaaaaaaag aagacacaga tcgcaaatgc acaggacagg agagaattgg aaaagattcc	1200
aactatgagc agggaggtaa agtccagttc gtgattgacg cagtctatgc tatggctcac	1260
gcccttcacc acatgaacaa ggatctctgt gctgactacc ggggtgtctg ccagagatg	1320
gagcaagctg gaggcaagaa gttgctgaag tatatacgca atgttaattt caatggtagt	1380
gctggcactc cagtgatgtt taacaagaac ggggatgcac ctgggcgtta tgacatcttt	1440
cagtaccaga ccacaaacac cagcaacccg ggttaccgtc tgatcgggca gtggacagac	1500
gaacttcagc tcaatataga agacatgcag tggggtaaag gagtccgaga gatacccgcc	1560
tcagtgtgca cactaccatg taagccagga cagagaaaga agacacagaa aggaactcct	1620
tgctgttggc cctgtgagcc ttgcgatggt taccagtac cgtttgatga gatgacatgc	1680
cagcattgcc cctatgacca gagggccaat gaaaatcgaa ccggatgcca ggatattccc	1740


```

atcatcaaac tggagtggca ctccccctcg gctgtgattc ctgtcttcct ggcaatgttg 1800
gggatcattg ccaccatctt tgtcatggcc actttcatcc gctacaatga cacgcccatt 1860
gtccgggcat ctgggcggga actcagctat gttcttttga cgggcatctt tctttgctac 1920
atcatcactt tctgatgat tgccaaacca gatgtggcag tgtgttcttt ccggcgagtt 1980
ttcttgggct tgggtatgtg catcagttat gcagccctct tgacgaaaac aaatcggatt 2040
tatcgcatat ttgagcaggg caagaaatca gtaacagctc ccagactcat aagcccaaca 2100
tcacaactgg caatcacttc cagtttaata tcagttcagc ttctaggggt gttcatttgg 2160
tttggtgttg atccacccaa catcatcata gactacgatg aacacaagac aatgaaçcct 2220
gagcaagcca gaggggttct caagtgtgac attacagatc tccaaatcat ttgctccttg 2280
ggatatagca ttcttctcat ggtcacatgt actgtgtatg ccatcaagac tcgggggtgta 2340
cccgagaatt ttaacgaagc caagcccatt ggattcacta tgtacacgac atgtatagta 2400
tggtttgcct tcattccaat tttttttggc accgctcaat cagcggaata gctctacata 2460
caaactacca cgcttacaat ctccatgaac ctaagtgcac cagtggcgct ggggatgcta 2520
tacatgccga aagtgtacat catcattttc caccctgaac tcaatgtcca gaaacggaag 2580
cgaagcttca aggcggtagt cacagcagcc accatgtcat cgaggctgtc acacaaaccc 2640
agtgacagac ccaacggtga ggcaaagacc gagctctgtg aaaacgtaga cccaaacagc 2700
cctgctgcaa aaaagaagta tgtcagttat aataacctgg ttatctaa 2748

```

<210> 67
 <211> 915
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 67

```

Met Val Gln Leu Arg Lys Leu Leu Arg Val Leu Thr Leu Met Lys Phe
1             5             10             15

```

```

Pro Cys Cys Val Leu Glu Val Leu Leu Cys Ala Leu Ala Ala Ala Ala
20             25             30

```

```

Arg Gly Gln Glu Met Tyr Ala Pro His Ser Ile Arg Ile Glu Gly Asp
35             40             45

```

Val Thr Leu Gly Gly Leu Phe Pro Val His Ala Lys Gly Pro Ser Gly
 50 55 60

Val Pro Cys Gly Asp Ile Lys Arg Glu Asn Gly Ile His Arg Leu Glu
 65 70 75 80

Ala Met Leu Tyr Ala Leu Asp Gln Ile Asn Ser Asp Pro Asn Leu Leu
 85 90 95

Pro Asn Val Thr Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp
 100 105 110

Thr Tyr Ala Leu Glu Gln Ser Leu Thr Phe Val Gln Ala Leu Ile Gln
 115 120 125

Lys Asp Thr Ser Asp Val Arg Cys Thr Asn Gly Glu Pro Pro Val Phe
 130 135 140

Val Lys Pro Glu Lys Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser
 145 150 155 160

Val Ser Ile Met Val Ala Asn Ile Leu Arg Leu Phe Gln Ile Pro Gln
 165 170 175

Ile Ser Tyr Ala Ser Thr Ala Pro Glu Leu Ser Asp Asp Arg Arg Tyr
 180 185 190

Asp Phe Phe Ser Arg Val Val Pro Pro Asp Ser Phe Gln Ala Gln Ala
 195 200 205

Met Val Asp Ile Val Lys Ala Leu Gly Trp Asn Tyr Val Ser Thr Leu
 210 215 220

Ala Ser Glu Gly Ser Tyr Gly Glu Lys Gly Val Glu Ser Phe Thr Gln
 225 230 235 240

Ile Ser Lys Glu Ala Gly Gly Leu Cys Ile Ala Gln Ser Val Arg Ile
 245 250 255

Pro Gln Glu Arg Lys Asp Arg Thr Ile Asp Phe Asp Arg Ile Ile Lys
 260 265 270

Gln Leu Leu Asp Thr Pro Asn Ser Arg Ala Val Val Ile Phe Ala Asn

275	280	285
Asp Glu Asp Ile Lys Gln Ile Leu Ala Ala Ala Lys Arg Ala Asp Gln 290 295 300		
Val Gly His Phe Leu Trp Val Gly Ser Asp Ser Trp Gly Ser Lys Ile 305 310 315 320		
Asn Pro Leu His Gln His Glu Asp Ile Ala Glu Gly Ala Ile Thr Ile 325 330 335		
Gln Pro Lys Arg Ala Thr Val Glu Gly Phe Asp Ala Tyr Phe Thr Ser 340 345 350		
Arg Thr Leu Glu Asn Asn Arg Arg Asn Val Trp Phe Ala Glu Tyr Trp 355 360 365		
Glu Glu Asn Phe Asn Cys Lys Leu Thr Ile Ser Gly Ser Lys Lys Glu 370 375 380		
Asp Thr Asp Arg Lys Cys Thr Gly Gln Glu Arg Ile Gly Lys Asp Ser 385 390 395 400		
Asn Tyr Glu Gln Glu Gly Lys Val Gln Phe Val Ile Asp Ala Val Tyr 405 410 415		
Ala Met Ala His Ala Leu His His Met Asn Lys Asp Leu Cys Ala Asp 420 425 430		
Tyr Arg Gly Val Cys Pro Glu Met Glu Gln Ala Gly Gly Lys Lys Leu 435 440 445		
Leu Lys Tyr Ile Arg Asn Val Asn Phe Asn Gly Ser Ala Gly Thr Pro 450 455 460		
Val Met Phe Asn Lys Asn Gly Asp Ala Pro Gly Arg Tyr Asp Ile Phe 465 470 475 480		
Gln Tyr Gln Thr Thr Asn Thr Ser Asn Pro Gly Tyr Arg Leu Ile Gly 485 490 495		
Gln Trp Thr Asp Glu Leu Gln Leu Asn Ile Glu Asp Met Gln Trp Gly 500 505 510		

Lys Gly Val Arg Glu Ile Pro Ala Ser Val Cys Thr Leu Pro Cys Lys
 515 520 525

Pro Gly Gln Arg Lys Lys Thr Gln Lys Gly Thr Pro Cys Cys Trp Thr
 530 535 540

Cys Glu Pro Cys Asp Gly Tyr Gln Tyr Gln Phe Asp Glu Met Thr Cys
 545 550 555 560

Gln His Cys Pro Tyr Asp Gln Arg Pro Asn Glu Asn Arg Thr Gly Cys
 565 570 575

Gln Asp Ile Pro Ile Ile Lys Leu Glu Trp His Ser Pro Ser Ala Val
 580 585 590

Ile Pro Val Phe Leu Ala Met Leu Gly Ile Ile Ala Thr Ile Phe Val
 595 600 605

Met Ala Thr Phe Ile Arg Tyr Asn Asp Thr Pro Ile Val Arg Ala Ser
 610 615 620

Gly Arg Glu Leu Ser Tyr Val Leu Leu Thr Gly Ile Phe Leu Cys Tyr
 625 630 635 640

Ile Ile Thr Phe Leu Met Ile Ala Lys Pro Asp Val Ala Val Cys Ser
 645 650 655

Phe Arg Arg Val Phe Leu Gly Leu Gly Met Cys Ile Ser Tyr Ala Ala
 660 665 670

Leu Leu Thr Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys
 675 680 685

Lys Ser Val Thr Ala Pro Arg Leu Ile Ser Pro Thr Ser Gln Leu Ala
 690 695 700

Ile Thr Ser Ser Leu Ile Ser Val Gln Leu Leu Gly Val Phe Ile Trp
 705 710 715 720

Phe Gly Val Asp Pro Pro Asn Ile Ile Ile Asp Tyr Asp Glu His Lys
 725 730 735

Thr Met Asn Pro Glu Gln Ala Arg Gly Val Leu Lys Cys Asp Ile Thr
 740 745 750

Asp Leu Gln Ile Ile Cys Ser Leu Gly Tyr Ser Ile Leu Leu Met Val
 755 760 765

Thr Cys Thr Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Asn Phe
 770 775 780

Asn Glu Ala Lys Pro Ile Gly Phe Thr Met Tyr Thr Thr Cys Ile Val
 785 790 795 800

Trp Leu Ala Phe Ile Pro Ile Phe Phe Gly Thr Ala Gln Ser Ala Glu
 805 810 815

Lys Leu Tyr Ile Gln Thr Thr Thr Leu Thr Ile Ser Met Asn Leu Ser
 820 825 830

Ala Ser Val Ala Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile
 835 840 845

Ile Phe His Pro Glu Leu Asn Val Gln Lys Arg Lys Arg Ser Phe Lys
 850 855 860

Ala Val Val Thr Ala Ala Thr Met Ser Ser Arg Leu Ser His Lys Pro
 865 870 875 880

Ser Asp Arg Pro Asn Gly Glu Ala Lys Thr Glu Leu Cys Glu Asn Val
 885 890 895

Asp Pro Asn Ser Pro Ala Ala Lys Lys Lys Tyr Val Ser Tyr Asn Asn
 900 905 910

Leu Val Ile
 915

<210> 68
 <211> 2748
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 68
 atggtccagc tgaggaagct gctccgcgtc ctgactttga tgaagttccc ctgctgcgtg 60
 ctggaggtgc tcctgtgcgc gctggcgcgc gcggcgcgc gccaggagat gtacgccccg 120
 cactcaatcc ggatcgaggg ggacgtcacc ctcggggggc tgttccccgt gcacgccaag 180
 ggtccagcg gagtgccctg cggcgacatc aagagggaaa acgggatcca caggctggaa 240
 gcgatgtct acgccctgga ccagatcaac agtgatccca acctactgcc caacgtgacg 300
 ctgggcgcgc ggatcctgga cacttgttcc agggacactt acgcgctcga acagtcgctt 360
 actttcgtcc aggcgctcat ccagaaggac acctccgacg tgcgctgcac caacggcgaa 420
 ccgccggttt tcgtcaagcc ggagaaagta gttggagtga ttggggcttc ggggagttcg 480
 gtctccatca tggtagccaa catcctgagg ctcttcaga tccccagat tagttatgca 540
 tcaacggcac ccgagctaag tgatgaccgg cgctatgact tcttctctcg cgtggtgcc 600
 cccgattcct tccaagccca ggccattgta gacattgtaa aggccctagg ctggaattat 660
 gtgtctaccc tcgcâtcgga aggaagtat ggagagaaag gtgtggagtc cttcacgcag 720
 atttccaaag aggcaggtgg actctgcatt gccagtcgc tgagaatccc ccaggaacgc 780
 aaagacagga ccattgactt tgatagaatt atcaaacagc tcttgacac ccccaactcc 840
 agggccgctc tgatttttgc caacgatgag gatataaagc agatccttgc agcagccaaa 900
 agagctgacc aagtggcca ttttcttgg gtgggatcag acagctgggg atccaaaata 960
 aaccactgc accagcatga agatâtcga gaaggggcca tcaccattca gcccaagcga 1020
 gccacggtgg aagggtttga tgcctacttt acgtcccga cacttgaaaa caacagaaga 1080
 aatgtatggt ttgccgaata ctgggaggaa aacttcaact gcaagttgac gattagtggg 1140
 tcaaaaaaag aagacacaga tcgcaaâtc acaggacagg agagaattgg aaaagattcc 1200
 aactatgagc aggagggtaa agtccagttc gtgattgacg cagtctatgc tatggctcac 1260
 gcccttcacc acatgaacaa ggatctctgt gctgactacc ggggtgtctg ccagagatg 1320
 gagcaagctg gaggcaagaa gttgctgaag tatatacga atgttaattt caatggtagt 1380
 gctggcactc cagtgatgtt taacaagaac ggggatgcac ctgggcgta tgacatcttt 1440
 cagtaccaga ccacaaacac cagcaaccgc ggttaccgtc tgatcgggca gtggacagac 1500
 gaacttcagc tcaatataga agacatgcag tggggtaaag gâtccgaga gatacccgcc 1560
 tcagtgtgca cactaccatg taagccagga cagagaaaga agacacagaa aggaactcct 1620
 tgctgttga cctgtgagcc ttgcgatggt taccagtacc agtttgatga gatgacatgc 1680

cagcattgcc cctatgacca gaggcccaat gaaaatcgaa ccggatgccca ggatattccc 1740
 atcatcaaac tggagtggca ctccccctgg gctgtgattc ctgtcttcct ggcaatgttg 1800
 gggatcattg ccaccatctt tgtcatggcc actttcatcc gctacaatga cagccccatt 1860
 gtccggggcat ctgggcgggg actcagctat gttcttttga cgggcatctt tctttgctac 1920
 atcatcactt tcctgatgat tgccaaacca gatgtggcag tgtgttcttt ccggcatgtt 1980
 ttcttgggct tgggtatgtg catcagttat gcagccctct tgacgaaaac aaatcggatt 2040
 tatcgcatat ttgagcaggg caagaaatca gtaacagctc ccagactcat aagcccaaca 2100
 tcacaactgg caatcacttc cagtttaata tcagttcagc ttctaggggt gttcatttgg 2160
 tttggtgttg atccacccaa catcatcata gactacgatg aacacaagac aatgaaccct 2220
 gagcaagcca gaggggttct caagtgtgac attacagatc tccaaatcat ttgctccttg 2280
 ggatatagca ttcttctcat ggtcacatgt actgtgtatg ccatcaagac tcgggggtga 2340
 cccgagaatt ttaacgaagc caagcccatt ggattcacta tgtacacgac atgtatagta 2400
 tggcttgctt tcattccaat ttttttggc accgctcaat cagcggaaaa gctctacata 2460
 caaactacca cgcttacaat ctccatgaac ctaagtgcac cagtggcgct ggggatgcta 2520
 tacatgccga aagtgtacat catcattttc caccctgaac tcaatgtcca gaaacggaag 2580
 cgaagcttca aggcggtagt cacagcagcc accatgtcat cgaggctgtc acacaaacct 2640
 agtgacagac ccaacggtga ggcaaagacc gagctctgtg aaaacgtaga cccaacagc 2700
 cctgctgcaa aaaagaagta tgtcagttat aataacctgg ttatctaa 2748

<210> 69
 <211> 915
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 69

Met Val Gln Leu Arg Lys Leu Leu Arg Val Leu Thr Leu Met Lys Phe
 1 5 10 15

Pro Cys Cys Val Leu Glu Val Leu Leu Cys Ala Leu Ala Ala Ala Ala
 20 25 30

Arg Gly Gln Glu Met Tyr Ala Pro His Ser Ile Arg Ile Glu Gly Asp
 35 40 45

Val Thr Leu Gly Gly Leu Phe Pro Val His Ala Lys Gly Pro Ser Gly
 50 55 60

Val Pro Cys Gly Asp Ile Lys Arg Glu Asn Gly Ile His Arg Leu Glu
 65 70 75 80

Ala Met Leu Tyr Ala Leu Asp Gln Ile Asn Ser Asp Pro Asn Leu Leu
 85 90 95

Pro Asn Val Thr Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp
 100 105 110

Thr Tyr Ala Leu Glu Gln Ser Leu Thr Phe Val Gln Ala Leu Ile Gln
 115 120 125

Lys Asp Thr Ser Asp Val Arg Cys Thr Asn Gly Glu Pro Pro Val Phe
 130 135 140

Val Lys Pro Glu Lys Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser
 145 150 155 160

Val Ser Ile Met Val Ala Asn Ile Leu Arg Leu Phe Gln Ile Pro Gln
 165 170 175

Ile Ser Tyr Ala Ser Thr Ala Pro Glu Leu Ser Asp Asp Arg Arg Tyr
 180 185 190

Asp Phe Phe Ser Arg Val Val Pro Pro Asp Ser Phe Gln Ala Gln Ala
 195 200 205

Met Val Asp Ile Val Lys Ala Leu Gly Trp Asn Tyr Val Ser Thr Leu
 210 215 220

Ala Ser Glu Gly Ser Tyr Gly Glu Lys Gly Val Glu Ser Phe Thr Gln
 225 230 235 240

Ile Ser Lys Glu Ala Gly Gly Leu Cys Ile Ala Gln Ser Val Arg Ile
 245 250 255

Pro Gln Glu Arg Lys Asp Arg Thr Ile Asp Phe Asp Arg Ile Ile Lys
 260 265 270

Gln Leu Leu Asp Thr Pro Asn Ser Arg Ala Val Val Ile Phe Ala Asn
 275 280 285

Asp Glu Asp Ile Lys Gln Ile Leu Ala Ala Ala Lys Arg Ala Asp Gln
 290 295 300

Val Gly His Phe Leu Trp Val Gly Ser Asp Ser Trp Gly Ser Lys Ile
 305 310 315 320

Asn Pro Leu His Gln His Glu Asp Ile Ala Glu Gly Ala Ile Thr Ile
 325 330 335

Gln Pro Lys Arg Ala Thr Val Glu Gly Phe Asp Ala Tyr Phe Thr Ser
 340 345 350

Arg Thr Leu Glu Asn Asn Arg Arg Asn Val Trp Phe Ala Glu Tyr Trp
 355 360 365

Glu Glu Asn Phe Asn Cys Lys Leu Thr Ile Ser Gly Ser Lys Lys Glu
 370 375 380

Asp Thr Asp Arg Lys Cys Thr Gly Gln Glu Arg Ile Gly Lys Asp Ser
 385 390 395 400

Asn Tyr Glu Gln Glu Gly Lys Val Gln Phe Val Ile Asp Ala Val Tyr
 405 410 415

Ala Met Ala His Ala Leu His His Met Asn Lys Asp Leu Cys Ala Asp
 420 425 430

Tyr Arg Gly Val Cys Pro Glu Met Glu Gln Ala Gly Gly Lys Lys Leu
 435 440 445

Leu Lys Tyr Ile Arg Asn Val Asn Phe Asn Gly Ser Ala Gly Thr Pro
 450 455 460

Val Met Phe Asn Lys Asn Gly Asp Ala Pro Gly Arg Tyr Asp Ile Phe
 465 470 475 480

Gln Tyr Gln Thr Thr Asn Thr Ser Asn Pro Gly Tyr Arg Leu Ile Gly
 485 490 495

Gln Trp Thr Asp Glu Leu Gln Leu Asn Ile Glu Asp Met Gln Trp Gly
 500 505 510

Lys Gly Val Arg Glu Ile Pro Ala Ser Val Cys Thr Leu Pro Cys Lys
 515 520 525

Pro Gly Gln Arg Lys Lys Thr Gln Lys Gly Thr Pro Cys Cys Trp Thr
 530 535 540

Cys Glu Pro Cys Asp Gly Tyr Gln Tyr Gln Phe Asp Glu Met Thr Cys
 545 550 555 560

Gln His Cys Pro Tyr Asp Gln Arg Pro Asn Glu Asn Arg Thr Gly Cys
 565 570 575

Gln Asp Ile Pro Ile Ile Lys Leu Glu Trp His Ser Pro Trp Ala Val
 580 585 590

Ile Pro Val Phe Leu Ala Met Leu Gly Ile Ile Ala Thr Ile Phe Val
 595 600 605

Met Ala Thr Phe Ile Arg Tyr Asn Asp Thr Pro Ile Val Arg Ala Ser
 610 615 620

Gly Arg Glu Leu Ser Tyr Val Leu Leu Thr Gly Ile Phe Leu Cys Tyr
 625 630 635 640

Ile Ile Thr Phe Leu Met Ile Ala Lys Pro Asp Val Ala Val Cys Ser
 645 650 655

Phe Arg His Val Phe Leu Gly Leu Gly Met Cys Ile Ser Tyr Ala Ala
 660 665 670

Leu Leu Thr Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys
 675 680 685

Lys Ser Val Thr Ala Pro Arg Leu Ile Ser Pro Thr Ser Gln Leu Ala
 690 695 700

Ile Thr Ser Ser Leu Ile Ser Val Gln Leu Leu Gly Val Phe Ile Trp
 705 710 715 720

Phe Gly Val Asp Pro Pro Asn Ile Ile Ile Asp Tyr Asp Glu His Lys

725

730

735

Thr Met Asn Pro Glu Gln Ala Arg Gly Val Leu Lys Cys Asp Ile Thr
 740 745 750

Asp Leu Gln Ile Ile Cys Ser Leu Gly Tyr Ser Ile Leu Leu Met Val
 755 760 765

Thr Cys Thr Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Asn Phe
 770 775 780

Asn Glu Ala Lys Pro Ile Gly Phe Thr Met Tyr Thr Thr Cys Ile Val
 785 790 795 800

Trp Leu Ala Phe Ile Pro Ile Phe Phe Gly Thr Ala Gln Ser Ala Glu
 805 810 815

Lys Leu Tyr Ile Gln Thr Thr Thr Leu Thr Ile Ser Met Asn Leu Ser
 820 825 830

Ala Ser Val Ala Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile
 835 840 845

Ile Phe His Pro Glu Leu Asn Val Gln Lys Arg Lys Arg Ser Phe Lys
 850 855 860

Ala Val Val Thr Ala Ala Thr Met Ser Ser Arg Leu Ser His Lys Pro
 865 870 875 880

Ser Asp Arg Pro Asn Gly Glu Ala Lys Thr Glu Leu Cys Glu Asn Val
 885 890 895

Asp Pro Asn Ser Pro Ala Ala Lys Lys Lys Tyr Val Ser Tyr Asn Asn
 900 905 910

Leu Val Ile
 915

<210> 70
 <211> 2748
 <212> DNA
 <213> Unknown

<220>

<223> Novel Sequence

<400> 70
 atgggtccagc tgaggaagct gctccgcgtc ctgactttga tgaagttccc ctgctgctgtg 60
 ctggagggtgc tcctgtgcgc gctggcgggcg gcggcgcgcg gccaggagat gtacgccccg 120
 cactcaatcc ggatcgaggg ggacgtcacc ctcggggggc tgttccccgt gcacgccaag 180
 ggtcccagcg gagtgccttg cggcgacatc aagagggaaa acgggatcca caggctggaa 240
 gcgatgtctt acgccctgga ccagatcaac agtgatccca acctactgcc caacgtgacg 300
 ctgggcgcgc ggatcctgga cacttgttcc agggacactt acgcgtcga acagtgcgtt 360
 actttcgtcc aggcgtcat ccagaaggac acctccgacg tgcgctgcac caacggcgaa 420
 ccgccggttt tcgtcaagcc ggagaaagta gttggagtga ttggggcttc ggggagttcg 480
 gtctccatca tggtagcaa catcctgagg ctcttcaga tccccagat tagttatgca 540
 tcaacggcac ccgagctaag tgatgaccgg cgctatgact tcttctctcg cgtggtgcca 600
 cccgattcct tccaagccca ggcatggtg gacattgtaa aggccctagg ctggaattat 660
 gtgtctaccc tcgcatcgga aggaagtat ggagagaaa gtgtggagtc cttcacgcag 720
 atttccaaag aggcaggtgg actctgcatt gccagtcgg tgagaatccc ccaggaacgc 780
 aaagacagga ccattgactt tgatagaatt atcaaacagc tcctggacac cccaactcc 840
 agggccgtcg tgatttttgc caacgatgag gatataaagc agatccttgc agcagccaaa 900
 agagctgacc aagtggcca ttttctttgg gtgggatcag acagctgggg atccaaaata 960
 aaccactgc accagcatga agatatcgca gaaggggcca tcaccattca gcccaagcga 1020
 gccacggtgg aagggtttga tgcctacttt acgtcccgt cacttgaaaa caacagaaga 1080
 aatgtatggt ttgccgaata ctgggaggaa aacttcaact gcaagttgac gattagtggg 1140
 tcaaaaaaag aagacacaga tcgcaaatgc acaggacagg agagaattgg aaaagattcc 1200
 aactatgagc aggagggtaa agtccagttc gtgattgacg cagtctatgc tatggctcac 1260
 gcccttcacc acatgaacaa ggatctctgt gctgactacc ggggtgtctg ccagagatg 1320
 gagcaagctg gaggaagaa gttgctgaag tatatacgca atgttaattt caatggtagt 1380
 gctggcactc cagtgatgtt taacaagaac ggggatgcac ctgggcgtta tgacatcttt 1440
 cagtaccaga ccacaaacac cagcaacccg gggtaccgtc tgatcgggca gtggacagac 1500
 gaacttcagc tcaatataga agacatgcag tggggtaaag gagtccgaga gatacccgcc 1560
 tcagtgtgca cactaccatg taagccagga cagagaaaga agacacagaa aggaactcct 1620

```

tgctgttgga cctgtgagcc ttgcgatggt taccagtacc agtttgatga gatgacatgc 1680
cagcattgcc cctatgacca gagggccaat gaaaatcgaa ccggatgcca ggatattccc 1740
atcatcaaac tggagtggca ctccccctgg gctgtgattc ctgtcttcct ggcaatgttg 1800
gggatcattg ccaccatctt tgtcatggcc actttcatcc gctacaatga cacgcccatt 1860
gtccgggcat ctgggcgga actcagctat gttcttttga cgggcatctt tctttgctac 1920
atcatcactt tcctgatgat tgccaaacca gatgtggcag tgtgttcttt ccggcgagtt 1980
ttcttgggct tgggtatgtg catcagttat gcagccctct tgacgaaaac aaatcggatt 2040
tatcgcatat ttgagcaggg caagaaatca gtaacagctc ccagactcat aagcccaaca 2100
tcacaactgg caatcacttc cagtttaata tcagttcagc ttctaggggt gttcatttgg 2160
tttgggtgtg atccacccaa catcatcata gactacgatg aacacaagac aatgaaccct 2220
gagcaagcca gaggggttct caagtgtgac attacagatc tccaaatcat ttgctccttg 2280
ggatatagca ttcttctcat ggtcacatgt tgtgtgtatg ccatcaagac tcgggggtgta 2340
cccgagaatt ttaacgaagc caagcccatt ggattcacta tgtacacgac atgtatagta 2400
tggcttgcct tcattccaat ttttttggc accgctcaat cagcggaaaa gctctacata 2460
caaaactacca cgcttacaat ctccatgaac ctaagtgcac cagtggcgct ggggatgcta 2520
tacatgccga aagtgtacat catcattttc caccctgaac tcaatgtcca gaaacggaag 2580
cgaagcttca aggcggtagt cacagcagcc accatgtcat cgaggctgtc acacaaaccc 2640
agtacagac ccaacggtga ggcaaagacc gagctctgtg aaaacgtaga cccaaacagc 2700
cctgctgcaa aaaagaagta tgtcagttat aataacctgg ttatctaa 2748

```

<210> 71
 <211> 915
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 71

```

Met Val Gln Leu Arg Lys Leu Leu Arg Val Leu Thr Leu Met Lys Phe
1           5           10           15

```

```

Pro Cys Cys Val Leu Glu Val Leu Leu Cys Ala Leu Ala Ala Ala Ala
20           25           30

```

Arg Gly Gln Glu Met Tyr Ala Pro His Ser Ile Arg Ile Glu Gly Asp
 35 40 45

Val Thr Leu Gly Gly Leu Phe Pro Val His Ala Lys Gly Pro Ser Gly
 50 55 60

Val Pro Cys Gly Asp Ile Lys Arg Glu Asn Gly Ile His Arg Leu Glu
 65 70 75 80

Ala Met Leu Tyr Ala Leu Asp Gln Ile Asn Ser Asp Pro Asn Leu Leu
 85 90 95

Pro Asn Val Thr Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp
 100 105 110

Thr Tyr Ala Leu Glu Gln Ser Leu Thr Phe Val Gln Ala Leu Ile Gln
 115 120 125

Lys Asp Thr Ser Asp Val Arg Cys Thr Asn Gly Glu Pro Pro Val Phe
 130 135 140

Val Lys Pro Glu Lys Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser
 145 150 155 160

Val Ser Ile Met Val Ala Asn Ile Leu Arg Leu Phe Gln Ile Pro Gln
 165 170 175

Ile Ser Tyr Ala Ser Thr Ala Pro Glu Leu Ser Asp Asp Arg Arg Tyr
 180 185 190

Asp Phe Phe Ser Arg Val Val Pro Pro Asp Ser Phe Gln Ala Gln Ala
 195 200 205

Met Val Asp Ile Val Lys Ala Leu Gly Trp Asn Tyr Val Ser Thr Leu
 210 215 220

Ala Ser Glu Gly Ser Tyr Gly Glu Lys Gly Val Glu Ser Phe Thr Gln
 225 230 235 240

Ile Ser Lys Glu Ala Gly Gly Leu Cys Ile Ala Gln Ser Val Arg Ile
 245 250 255

Pro Gln Glu Arg Lys Asp Arg Thr Ile Asp Phe Asp Arg Ile Ile Lys

260	265	270
Gln Leu Leu Asp Thr Pro Asn Ser Arg Ala Val Val Ile Phe Ala Asn 275 280 285		
Asp Glu Asp Ile Lys Gln Ile Leu Ala Ala Ala Lys Arg Ala Asp Gln 290 295 300		
Val Gly His Phe Leu Trp Val Gly Ser Asp Ser Trp Gly Ser Lys Ile 305 310 315 320		
Asn Pro Leu His Gln His Glu Asp Ile Ala Glu Gly Ala Ile Thr Ile 325 330 335		
Gln Pro Lys Arg Ala Thr Val Glu Gly Phe Asp Ala Tyr Phe Thr Ser 340 345 350		
Arg Thr Leu Glu Asn Asn Arg Arg Asn Val Trp Phe Ala Glu Tyr Trp 355 360 365		
Glu Glu Asn Phe Asn Cys Lys Leu Thr Ile Ser Gly Ser Lys Lys Glu 370 375 380		
Asp Thr Asp Arg Lys Cys Thr Gly Gln Glu Arg Ile Gly Lys Asp Ser 385 390 395 400		
Asn Tyr Glu Gln Glu Gly Lys Val Gln Phe Val Ile Asp Ala Val Tyr 405 410 415		
Ala Met Ala His Ala Leu His His Met Asn Lys Asp Leu Cys Ala Asp 420 425 430		
Tyr Arg Gly Val Cys Pro Glu Met Glu Gln Ala Gly Gly Lys Lys Leu 435 440 445		
Leu Lys Tyr Ile Arg Asn Val Asn Phe Asn Gly Ser Ala Gly Thr Pro 450 455 460		
Val Met Phe Asn Lys Asn Gly Asp Ala Pro Gly Arg Tyr Asp Ile Phe 465 470 475 480		
Gln Tyr Gln Thr Thr Asn Thr Ser Asn Pro Gly Tyr Arg Leu Ile Gly 485 490 495		

Gln Trp Thr Asp Glu Leu Gln Leu Asn Ile Glu Asp Met Gln Trp Gly
 500 505 510

Lys Gly Val Arg Glu Ile Pro Ala Ser Val Cys Thr Leu Pro Cys Lys
 515 520 525

Pro Gly Gln Arg Lys Lys Thr Gln Lys Gly Thr Pro Cys Cys Trp Thr
 530 535 540

Cys Glu Pro Cys Asp Gly Tyr Gln Tyr Gln Phe Asp Glu Met Thr Cys
 545 550 555 560

Gln His Cys Pro Tyr Asp Gln Arg Pro Asn Glu Asn Arg Thr Gly Cys
 565 570 575

Gln Asp Ile Pro Ile Ile Lys Leu Glu Trp His Ser Pro Trp Ala Val
 580 585 590

Ile Pro Val Phe Leu Ala Met Leu Gly Ile Ile Ala Thr Ile Phe Val
 595 600 605

Met Ala Thr Phe Ile Arg Tyr Asn Asp Thr Pro Ile Val Arg Ala Ser
 610 615 620

Gly Arg Glu Leu Ser Tyr Val Leu Leu Thr Gly Ile Phe Leu Cys Tyr
 625 630 635 640

Ile Ile Thr Phe Leu Met Ile Ala Lys Pro Asp Val Ala Val Cys Ser
 645 650 655

Phe Arg Arg Val Phe Leu Gly Leu Gly Met Cys Ile Ser Tyr Ala Ala
 660 665 670

Leu Leu Thr Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys
 675 680 685

Lys Ser Val Thr Ala Pro Arg Leu Ile Ser Pro Thr Ser Gln Leu Ala
 690 695 700

Ile Thr Ser Ser Leu Ile Ser Val Gln Leu Leu Gly Val Phe Ile Trp
 705 710 715 720

Phe Gly Val Asp Pro Pro Asn Ile Ile Ile Asp Tyr Asp Glu His Lys
725 730 735

Thr Met Asn Pro Glu Gln Ala Arg Gly Val Leu Lys Cys Asp Ile Thr
740 745 750

Asp Leu Gln Ile Ile Cys Ser Leu Gly Tyr Ser Ile Leu Leu Met Val
755 760 765

Thr Cys Cys Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Asn Phe
770 775 780

Asn Glu Ala Lys Pro Ile Gly Phe Thr Met Tyr Thr Thr Cys Ile Val
785 790 795 800

Trp Leu Ala Phe Ile Pro Ile Phe Phe Gly Thr Ala Gln Ser Ala Glu
805 810 815

Lys Leu Tyr Ile Gln Thr Thr Thr Leu Thr Ile Ser Met Asn Leu Ser
820 825 830

Ala Ser Val Ala Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile
835 840 845

Ile Phe His Pro Glu Leu Asn Val Gln Lys Arg Lys Arg Ser Phe Lys
850 855 860

Ala Val Val Thr Ala Ala Thr Met Ser Ser Arg Leu Ser His Lys Pro
865 870 875 880

Ser Asp Arg Pro Asn Gly Glu Ala Lys Thr Glu Leu Cys Glu Asn Val
885 890 895

Asp Pro Asn Ser Pro Ala Ala Lys Lys Lys Tyr Val Ser Tyr Asn Asn
900 905 910

Leu Val Ile
915

<210> 72
<211> 2748
<212> DNA
<213> Unknown

<220>

<223> Novel Sequence

<400> 72

atggtccagc tgaggaagct gctccgcgtc ctgactttga tgaagttccc ctgctgcgtg	60
ctggagggtgc tcctgtgcgc gctggcggcg gcggcgcgcg gccaggagat gtacgccccg	120
cactcaatcc ggatcgaggg ggacgtcacc ctcggggggc tgttccccgt gcacgccaag	180
gggtccagcg gagtgccttg cggcgacatc aagagggaaa acgggatcca caggctggaa	240
gcgatgctct acgccctgga ccagatcaac agtgatccca acctactgcc caacgtgacg	300
ctgggcgcgc ggatcctgga cacttggttc agggacactt acgcgctcga acagtgcgtt	360
actttcgtcc aggcgctcat ccagaaggac acctccgacg tgcgctgcac caacggcgaa	420
ccgccggttt tcgtcaagcc ggagaaagta gttggagtga ttggggcttc ggggagttcg	480
gtctccatca tggtagccaa catcctgagg ctcttccaga tccccagat tagttatgca	540
tcaacggcac ccgagctaag tgatgaccgg cgctatgact tcttctctcg cgtggtgcc	600
cccgattcct tccaagccca ggccatggtg gacattgtaa aggccctagg ctggaattat	660
gtgtctaccc tcgcatcgga aggaagttat ggagagaaag gtgtggagtc cttcacgcag	720
atttccaaag aggcaggtgg actctgcatt gcccagtcgg tgagaatccc ccaggaacgc	780
aaagacagga ccattgactt tgatagaatt atcaaacagc tcctggacac cccaactcc	840
agggccgtcg tgatttttgc caacgatgag gatataaagc agatccttgc agcagccaaa	900
agagctgacc aagttggcca ttttctttgg gtgggatcag acagctgggg atccaaaata	960
aaccactgc accagcatga agatatcgca gaagggggcca tcaccattca gcccaagcga	1020
gccacggtgg aagggtttga tgcctacttt acgtcccgtg cacttgaaaa caacagaaga	1080
aatgtatggt ttgccgaata ctgggaggaa aacttcaact gcaagttgac gattagtggg	1140
tcaaaaaaag aagacacaga tcgcaaatgc acaggacagg agagaattgg aaaagattcc	1200
aactatgagc agggaggtaa agtccagttc gtgattgacg cagtctatgc tatggctcac	1260
gcccttcacc acatgaacaa ggatctctgt gctgactacc ggggtgtctg ccagagatg	1320
gagcaagctg gaggcaagaa gttgctgaag tatatacgca atgttaattt caatggtagt	1380
gctggcactc cagtgatgtt taacaagaac ggggatgcac ctgggcgtta tgacatcttt	1440
cagtaccaga ccacaaacac cagcaacccg ggttaccgtc tgatcgggca gtggacagac	1500
gaacttcagc tcaatataga agacatgcag tggggtaaag ggtccgaga gatacccgcc	1560

```

tcagtgtgca cactaccatg taagccagga cagagaaaga agacacagaa aggaactcct 1620
tgctgttgga cctgtgagcc ttgcatggt taccagtacc agtttgatga gatgacatgc 1680
cagcattgcc cctatgacca gaggcccaat gaaaatcgaa ccggatgcca ggatattccc 1740
atcatcaaac tggagtggca ctccccctgg gctgtgattc ctgtcttcct ggcaatgttg 1800
gggatcattg ccaccatctt tgtcatggcc actttcatcc gctacaatga cagccccatt 1860
gtccgggcat ctgggcggga actcagctat gttcttttga cgggcatctt tctttgctac 1920
atcatcactt tctgatgat tgccaaacca gatgtggcag tgtgttcttt ccggcgagtt 1980
ttcttgggct tgggtatgtg catcagttat gcagccctct tgacgaaaac aaatcggatt 2040
tatcgcatat ttgagcaggg caagaaatca gtaacagctc ccagactcat aagcccaaca 2100
tcacaactgg caatcacttc cagtttaata tcagttcagc ttctaggggt gttcatttgg 2160
tttgggtgtg atccacccaa catcatcata gactacgatg aacacaagac aatgaaccct 2220
gagcaagcca gaggggttct caagtgtgac attacagatc tccaaatcat ttgctccttg 2280
ggatatagca ttcttctcat ggtcacatgt actgtgtatg ccatcaagac tcgggggtgta 2340
cccgagaatt ttaacgaagc caagcccaag ggattcacta tgtacacgac atgtatagta 2400
tggcttgcc tcatccaat ttttttggc accgctcaat cagcggaaaa gctctacata 2460
caaactacca cgcttacaat ctccatgaac ctaagtgcac cagtggcgct ggggatgcta 2520
tacatgccga aagtgtacat catcattttc caccctgaac tcaatgtcca gaaacggaag 2580
cgaagcttca aggcggtagt cacagcagcc accatgtcat cgaggctgtc acacaaaccc 2640
agtgacagac ccaacggtga ggcaaagacc gagctctgtg aaaacgtaga cccaaacagc 2700
cctgctgcaa aaaagaagta tgtcagttat aataacctgg ttatctaa 2748

```

<210> 73
 <211> 915
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 73

```

Met Val Gln Leu Arg Lys Leu Leu Arg Val Leu Thr Leu Met Lys Phe
1           5           10           15

```

```

Pro Cys Cys Val Leu Glu Val Leu Leu Cys Ala Leu Ala Ala Ala Ala
20           25           30

```

Arg Gly Gln Glu Met Tyr Ala Pro His Ser Ile Arg Ile Glu Gly Asp
 35 40 45

Val Thr Leu Gly Gly Leu Phe Pro Val His Ala Lys Gly Pro Ser Gly
 50 55 60

Val Pro Cys Gly Asp Ile Lys Arg Glu Asn Gly Ile His Arg Leu Glu
 65 70 75 80

Ala Met Leu Tyr Ala Leu Asp Gln Ile Asn Ser Asp Pro Asn Leu Leu
 85 90 95

Pro Asn Val Thr Leu Gly Ala Arg Ile Leu Asp Thr Cys Ser Arg Asp
 100 105 110

Thr Tyr Ala Leu Glu Gln Ser Leu Thr Phe Val Gln Ala Leu Ile Gln
 115 120 125

Lys Asp Thr Ser Asp Val Arg Cys Thr Asn Gly Glu Pro Pro Val Phe
 130 135 140

Val Lys Pro Glu Lys Val Val Gly Val Ile Gly Ala Ser Gly Ser Ser
 145 150 155 160

Val Ser Ile Met Val Ala Asn Ile Leu Arg Leu Phe Gln Ile Pro Gln
 165 170 175

Ile Ser Tyr Ala Ser Thr Ala Pro Glu Leu Ser Asp Asp Arg Arg Tyr
 180 185 190

Asp Phe Phe Ser Arg Val Val Pro Pro Asp Ser Phe Gln Ala Gln Ala
 195 200 205

Met Val Asp Ile Val Lys Ala Leu Gly Trp Asn Tyr Val Ser Thr Leu
 210 215 220

Ala Ser Glu Gly Ser Tyr Gly Glu Lys Gly Val Glu Ser Phe Thr Gln
 225 230 235 240

Ile Ser Lys Glu Ala Gly Gly Leu Cys Ile Ala Gln Ser Val Arg Ile
 245 250 255

Pro Gln Glu Arg Lys Asp Arg Thr Ile Asp Phe Asp Arg Ile Ile Lys
 260 265 270

Gln Leu Leu Asp Thr Pro Asn Ser Arg Ala Val Val Ile Phe Ala Asn
 275 280 285

Asp Glu Asp Ile Lys Gln Ile Leu Ala Ala Ala Lys Arg Ala Asp Gln
 290 295 300

Val Gly His Phe Leu Trp Val Gly Ser Asp Ser Trp Gly Ser Lys Ile
 305 310 315 320

Asn Pro Leu His Gln His Glu Asp Ile Ala Glu Gly Ala Ile Thr Ile
 325 330 335

Gln Pro Lys Arg Ala Thr Val Glu Gly Phe Asp Ala Tyr Phe Thr Ser
 340 345 350

Arg Thr Leu Glu Asn Asn Arg Arg Asn Val Trp Phe Ala Glu Tyr Trp
 355 360 365

Glu Glu Asn Phe Asn Cys Lys Leu Thr Ile Ser Gly Ser Lys Lys Glu
 370 375 380

Asp Thr Asp Arg Lys Cys Thr Gly Gln Glu Arg Ile Gly Lys Asp Ser
 385 390 395 400

Asn Tyr Glu Gln Glu Gly Lys Val Gln Phe Val Ile Asp Ala Val Tyr
 405 410 415

Ala Met Ala His Ala Leu His His Met Asn Lys Asp Leu Cys Ala Asp
 420 425 430

Tyr Arg Gly Val Cys Pro Glu Met Glu Gln Ala Gly Gly Lys Lys Leu
 435 440 445

Leu Lys Tyr Ile Arg Asn Val Asn Phe Asn Gly Ser Ala Gly Thr Pro
 450 455 460

Val Met Phe Asn Lys Asn Gly Asp Ala Pro Gly Arg Tyr Asp Ile Phe
 465 470 475 480

Gln Tyr Gln Thr Thr Asn Thr Ser Asn Pro Gly Tyr Arg Leu Ile Gly
485 490 495

Gln Trp Thr Asp Glu Leu Gln Leu Asn Ile Glu Asp Met Gln Trp Gly
500 505 510

Lys Gly Val Arg Glu Ile Pro Ala Ser Val Cys Thr Leu Pro Cys Lys
515 520 525

Pro Gly Gln Arg Lys Lys Thr Gln Lys Gly Thr Pro Cys Cys Trp Thr
530 535 540

Cys Glu Pro Cys Asp Gly Tyr Gln Tyr Gln Phe Asp Glu Met Thr Cys
545 550 555 560

Gln His Cys Pro Tyr Asp Gln Arg Pro Asn Glu Asn Arg Thr Gly Cys
565 570 575

Gln Asp Ile Pro Ile Ile Lys Leu Glu Trp His Ser Pro Trp Ala Val
580 585 590

Ile Pro Val Phe Leu Ala Met Leu Gly Ile Ile Ala Thr Ile Phe Val
595 600 605

Met Ala Thr Phe Ile Arg Tyr Asn Asp Thr Pro Ile Val Arg Ala Ser
610 615 620

Gly Arg Glu Leu Ser Tyr Val Leu Leu Thr Gly Ile Phe Leu Cys Tyr
625 630 635 640

Ile Ile Thr Phe Leu Met Ile Ala Lys Pro Asp Val Ala Val Cys Ser
645 650 655

Phe Arg Arg Val Phe Leu Gly Leu Gly Met Cys Ile Ser Tyr Ala Ala
660 665 670

Leu Leu Thr Lys Thr Asn Arg Ile Tyr Arg Ile Phe Glu Gln Gly Lys
675 680 685

Lys Ser Val Thr Ala Pro Arg Leu Ile Ser Pro Thr Ser Gln Leu Ala
690 695 700

Ile Thr Ser Ser Leu Ile Ser Val Gln Leu Leu Gly Val Phe Ile Trp

705	710	715	720
Phe Gly Val Asp Pro	Pro Asn Ile Ile	Ile Asp Tyr Asp Glu His Lys	
	725	730	735
Thr Met Asn Pro Glu Gln Ala Arg Gly Val Leu Lys Cys Asp Ile Thr			
	740	745	750
Asp Leu Gln Ile Ile Cys Ser Leu Gly Tyr Ser Ile Leu Leu Met Val			
	755	760	765
Thr Cys Thr Val Tyr Ala Ile Lys Thr Arg Gly Val Pro Glu Asn Phe			
	770	775	780
Asn Glu Ala Lys Pro Lys Gly Phe Thr Met Tyr Thr Thr Cys Ile Val			
	785	790	795
Trp Leu Ala Phe Ile Pro Ile Phe Phe Gly Thr Ala Gln Ser Ala Glu			
	805	810	815
Lys Leu Tyr Ile Gln Thr Thr Thr Leu Thr Ile Ser Met Asn Leu Ser			
	820	825	830
Ala Ser Val Ala Leu Gly Met Leu Tyr Met Pro Lys Val Tyr Ile Ile			
	835	840	845
Ile Phe His Pro Glu Leu Asn Val Gln Lys Arg Lys Arg Ser Phe Lys			
	850	855	860
Ala Val Val Thr Ala Ala Thr Met Ser Ser Arg Leu Ser His Lys Pro			
	865	870	875
Ser Asp Arg Pro Asn Gly Glu Ala Lys Thr Glu Leu Cys Glu Asn Val			
	885	890	895
Asp Pro Asn Ser Pro Ala Ala Lys Lys Lys Tyr Val Ser Tyr Asn Asn			
	900	905	910
Leu Val Ile			
	915		

<210> 74
 <211> 1842

<212> DNA

<213> Unknown

<220>

<223> Novel Sequence

<400> 74

```

atgcgagccc cgggcgcgct tctcgccgc atgtcgcggc tactgcttct gctactgctc      60
aaggtgtctg cctcttctgc cctcggggtc gccctgcgt ccagaaacga aacttgtctg      120
ggggagagct gtgcacctac agtgatccag cgccgcggca gggacgcctg gggaccggga      180
aattctgcaa gagacgttct gcgagccoga gacccaggg aggagcaggg ggcagcgttt      240
cttcggggac ctcctggga cctgccggcg gcccgggcc gtgaccggc tgcaggcaga      300
ggggcggagg cgtcggcagc cggacccccg ggacctcaa ccaggccacc tggcccctgg      360
aggtggaag gtgctcgggg tcaggagcct tctgaaactt tggggagagg gaacccacg      420
gccctcagc tcttccttca gatctcagag gaggaagaga agggcccag aggcgtggc      480
atttcgggc gtagccagga gcagagtgtg aagacagtcc ccggagccag cgatcttttt      540
tactggccaa ggagagccgg gaaactccag ggtcccacc acaagcccct gtccaagacg      600
gccaatggac tggcggggca cgaagggtgg acaattgcac tcccgggccc ggcgctggcc      660
cagaatggat ccttgggtga aggaatccat gagcctgggg gtccccgccc gggaaacagc      720
acgaaccggc gtgtgagact gaagaacccc ttctaccgc tgaccagga gtcctatgga      780
gcctacgcgg tcatgtgtct gtccgtggtg atcttcggga ccggcatcat tggcaacctg      840
gcggtgatgt gcatcgtgtg ccacaactac tacatgcgga gcatctcaa ctccctcttg      900
gccaacctgg cttctggga ctttctcacc atcttcttct gccttcgct ggtcatcttc      960
cacgagctga ccaagaagtg gctgctggag gacttctcct gcaagatcgt gccctatata     1020
gaggtogett ctctgggagt caccaccttc acccgatgtg ctctgtgcat agaccgcttc     1080
cgtgctgcca ccaacgtaca gatgtactac gaaatgatcg aaaactgttc ctcaacaact     1140
gccaaacttg ctgttatatg ggtgggagct ctattgttag cacttcaga agttgttctc     1200
cgccagctga gcaaggagga tttggggttt agtggccgag ctccggcaga aaggtgcatt     1260
attaagatct tcctgatatt accagacacc atctatgttc tagccctcac ctacgacagt     1320
gcgagactgt ggtggtatatt tggctgttac tttgtttgc ccacgctttt caccatcacc     1380
tgctctctag tgactgcgag gaaaatccgc aaagcagaga aagcctgtac ccgagggaat     1440
aaacggcaga ttcaactaga ggtcagatg aactgtacag tagtggcact gaccatttta     1500

```


tatggatttt gcattattcc tgaaaatata tgcaacattg ttactgccta catggctaca 1560
 ggggtttcac agcagacaat ggacctcctt aatatcatca gccagttcct tttgttcttt 1620
 aagtcctgtg tcaccccagt cctccttttc tgtctctgca aacccttcag tcgggccttc 1680
 atggagtgtc gctgctgttg ctgtgaggaa tgcattcaga agtcttcaac ggtgaccagt 1740
 gatgacaatg acaacgagta caccacggaa ctcgaactct cgcctttcag taccatacgc 1800
 cgtgaaatgt ccacttttgc ttctgtcgga actcattgct ga 1842

<210> 75
 <211> 613
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 75

Met Arg Ala Pro Gly Ala Leu Leu Ala Arg Met Ser Arg Leu Leu Leu
 1 5 10 15

Leu Leu Leu Leu Lys Val Ser Ala Ser Ser Ala Leu Gly Val Ala Pro
 20 25 30

Ala Ser Arg Asn Glu Thr Cys Leu Gly Glu Ser Cys Ala Pro Thr Val
 35 40 45

Ile Gln Arg Arg Gly Arg Asp Ala Trp Gly Pro Gly Asn Ser Ala Arg
 50 55 60

Asp Val Leu Arg Ala Arg Ala Pro Arg Glu Glu Gln Gly Ala Ala Phe
 65 70 75 80

Leu Ala Gly Pro Ser Trp Asp Leu Pro Ala Ala Pro Gly Arg Asp Pro
 85 90 95

Ala Ala Gly Arg Gly Ala Glu Ala Ser Ala Ala Gly Pro Pro Gly Pro
 100 105 110

Pro Thr Arg Pro Pro Gly Pro Trp Arg Trp Lys Gly Ala Arg Gly Gln
 115 120 125

Glu Pro Ser Glu Thr Leu Gly Arg Gly Asn Pro Thr Ala Leu Gln Leu
 130 135 140

Phe Leu Gln Ile Ser Glu Glu Glu Glu Lys Gly Pro Arg Gly Ala Gly
145 150 155 160

Ile Ser Gly Arg Ser Gln Glu Gln Ser Val Lys Thr Val Pro Gly Ala
165 170 175

Ser Asp Leu Phe Tyr Trp Pro Arg Arg Ala Gly Lys Leu Gln Gly Ser
180 185 190

His His Lys Pro Leu Ser Lys Thr Ala Asn Gly Leu Ala Gly His Glu
195 200 205

Gly Trp Thr Ile Ala Leu Pro Gly Arg Ala Leu Ala Gln Asn Gly Ser
210 215 220

Leu Gly Glu Gly Ile His Glu Pro Gly Gly Pro Arg Arg Gly Asn Ser
225 230 235 240

Thr Asn Arg Arg Val Arg Leu Lys Asn Pro Phe Tyr Pro Leu Thr Gln
245 250 255

Glu Ser Tyr Gly Ala Tyr Ala Val Met Cys Leu Ser Val Val Ile Phe
260 265 270

Gly Thr Gly Ile Ile Gly Asn Leu Ala Val Met Cys Ile Val Cys His
275 280 285

Asn Tyr Tyr Met Arg Ser Ile Ser Asn Ser Leu Leu Ala Asn Leu Ala
290 295 300

Phe Trp Asp Phe Leu Ile Ile Phe Phe Cys Leu Pro Leu Val Ile Phe
305 310 315 320

His Glu Leu Thr Lys Lys Trp Leu Leu Glu Asp Phe Ser Cys Lys Ile
325 330 335

Val Pro Tyr Ile Glu Val Ala Ser Leu Gly Val Thr Thr Phe Thr Arg
340 345 350

Cys Ala Leu Cys Ile Asp Arg Phe Arg Ala Ala Thr Asn Val Gln Met
355 360 365

Tyr Tyr Glu Met Ile Glu Asn Cys Ser Ser Thr Thr Ala Lys Leu Ala
 370 375 380

Val Ile Trp Val Gly Ala Leu Leu Leu Ala Leu Pro Glu Val Val Leu
 385 390 395 400

Arg Gln Leu Ser Lys Glu Asp Leu Gly Phe Ser Gly Arg Ala Pro Ala
 405 410 415

Glu Arg Cys Ile Ile Lys Ile Ser Pro Asp Leu Pro Asp Thr Ile Tyr
 420 425 430

Val Leu Ala Leu Thr Tyr Asp Ser Ala Arg Leu Trp Trp Tyr Phe Gly
 435 440 445

Cys Tyr Phe Cys Leu Pro Thr Leu Phe Thr Ile Thr Cys Ser Leu Val
 450 455 460

Thr Ala Arg Lys Ile Arg Lys Ala Glu Lys Ala Cys Thr Arg Gly Asn
 465 470 475 480

Lys Arg Gln Ile Gln Leu Glu Ser Gln Met Asn Cys Thr Val Val Ala
 485 490 495

Leu Thr Ile Leu Tyr Gly Phe Cys Ile Ile Pro Glu Asn Ile Cys Asn
 500 505 510

Ile Val Thr Ala Tyr Met Ala Thr Gly Val Ser Gln Gln Thr Met Asp
 515 520 525

Leu Leu Asn Ile Ile Ser Gln Phe Leu Leu Phe Phe Lys Ser Cys Val
 530 535 540

Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe
 545 550 555 560

Met Glu Cys Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser
 565 570 575

Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu
 580 585 590

Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser
 595 600 605

Val Gly Thr His Cys
 610

<210> 76

<211> 1842

<212> DNA

<213> Homo sapiens

<400> 76

```

atgcgagccc cgggcgcgct tctcgcccgc atgtcgcggc tactgcttct gctactgctc      60
aagggtgtctg cctcttctgc cctcggggtc gccctgcgt ccagaaacga aacttgctctg      120
ggggagagct gtgcacctac agtgatccag cgccgcggca gggacgcctg gggaccggga      180
aattctgcaa gagacgttct gcgagcccga gcaccaggg aggagcaggg ggcagcgttt      240
cttgcgggac cctcctggga cctgccggcg gccccgggcc gtgaccgggc tgcaggcaga      300
ggggcgagg cgtcggcagc cggacccccg ggacctcaa ccaggccacc tggcccctgg      360
aggtggaaag gtgctcgggg tcaggagcct tctgaaactt tggggagagg gaacccacg      420
gccctccagc tcttccttca gatctcagag gaggaagaga agggctccag aggcgctggc      480
atttcgggc gtagccagga gcagagtgtg aagacagtcc ccggagccag cgatcttttt      540
tactggccaa ggagagccgg gaaactccag ggttcccacc acaagcccct gtccaagacg      600
gccaatggac tggcggggca cgaagggtag acaattgcac tcccgggccg ggcgctggcc      660
cagaatggat ccttgggtga aggaatccat gagcctgggg gtccccgccg gggaaacagc      720
acgaaccggc gtgtgagact gaagaacccc ttctaccgc tgaccagga gtcctatgga      780
gcctacgcgg tcatgtgtct gtccgtggtg atcttcggga cggcatcat tggcaacctg      840
gcggtgatgt gcatcgtgtg ccacaactac tacatgcgga gcatctcaa ctccctcttg      900
gccaacctgg ccttctggga ctttctcctc atcttcttct gccttcgct ggtcatcttc      960
cacgagctga ccaagaagtg gctgctggag gacttctcct gcaagatcgt gccctatata     1020
gaggtcgctt ctctgggagt caccaccttc accttatgtg ctctgtgcat agaccgcttc     1080
cgtgctgcca ccaacgtaca gatgtactac gaaatgatcg aaaactgttc ctcaacaact     1140
gccaaacttg ctgttatatg ggtgggagct ctattgttag cacttcaga agttgttctc     1200
cgccagctga gcaaggagga tttggggttt agtggccgag ctccggcaga aagggtgcatt     1260
attaagatct ctccatgatt accagacacc atctatgttc tagccctcac ctacgacagt     1320

```

gcgagactgt ggtggtatatt tggctgttac ttttgtttgc ccacgctttt caccatcacc 1380
 tgctctctag tgactgcgag gaaaatccgc aaagcagaga aagcctgtac ccgaggggaat 1440
 aaacggcaga ttcaactaga gagtcagatg aactgtacag tagtggcact gaccatttta 1500
 tatggatttt gcattattcc tgaaaatata tgcaacattg ttactgccta catggctaca 1560
 ggggtttcac agcagacaat ggacctcctt aatatcatca gccagttcct tttgttcttt 1620
 aagtcctatg tcaccccgat cctccttttc tgtctctgca aacccttcag tcgggccttc 1680
 atggagtgtg gctgctgttg ctgtgaggaa tgcattcaga agtcttcaac ggtgaccagt 1740
 gatgacaatg acaacgagta caccacggaa ctgaactct cgcctttcag taccatacgc 1800
 cgtgaaatgt ccacttttgc ttctgtcgga actcattgct ga 1842

<210> 77
 <211> 613
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 77

Met Arg Ala Pro Gly Ala Leu Leu Ala Arg Met Ser Arg Leu Leu Leu
 1 5 10 15

Leu Leu Leu Leu Lys Val Ser Ala Ser Ser Ala Leu Gly Val Ala Pro
 20 25 30

Ala Ser Arg Asn Glu Thr Cys Leu Gly Glu Ser Cys Ala Pro Thr Val
 35 40 45

Ile Gln Arg Arg Gly Arg Asp Ala Trp Gly Pro Gly Asn Ser Ala Arg
 50 55 60

Asp Val Leu Arg Ala Arg Ala Pro Arg Glu Glu Gln Gly Ala Ala Phe
 65 70 75 80

Leu Ala Gly Pro Ser Trp Asp Leu Pro Ala Ala Pro Gly Arg Asp Pro
 85 90 95

Ala Ala Gly Arg Gly Ala Glu Ala Ser Ala Ala Gly Pro Pro Gly Pro
 100 105 110

Pro Thr Arg Pro Pro Gly Pro Trp Arg Trp Lys Gly Ala Arg Gly Gln
 115 120 125

Glu Pro Ser Glu Thr Leu Gly Arg Gly Asn Pro Thr Ala Leu Gln Leu
 130 135 140

Phe Leu Gln Ile Ser Glu Glu Glu Glu Lys Gly Pro Arg Gly Ala Gly
 145 150 155 160

Ile Ser Gly Arg Ser Gln Glu Gln Ser Val Lys Thr Val Pro Gly Ala
 165 170 175

Ser Asp Leu Phe Tyr Trp Pro Arg Arg Ala Gly Lys Leu Gln Gly Ser
 180 185 190

His His Lys Pro Leu Ser Lys Thr Ala Asn Gly Leu Ala Gly His Glu
 195 200 205

Gly Trp Thr Ile Ala Leu Pro Gly Arg Ala Leu Ala Gln Asn Gly Ser
 210 215 220

Leu Gly Glu Gly Ile His Glu Pro Gly Gly Pro Arg Arg Gly Asn Ser
 225 230 235 240

Thr Asn Arg Arg Val Arg Leu Lys Asn Pro Phe Tyr Pro Leu Thr Gln
 245 250 255

Glu Ser Tyr Gly Ala Tyr Ala Val Met Cys Leu Ser Val Val Ile Phe
 260 265 270

Gly Thr Gly Ile Ile Gly Asn Leu Ala Val Met Cys Ile Val Cys His
 275 280 285

Asn Tyr Tyr Met Arg Ser Ile Ser Asn Ser Leu Leu Ala Asn Leu Ala
 290 295 300

Phe Trp Asp Phe Leu Ile Ile Phe Phe Cys Leu Pro Leu Val Ile Phe
 305 310 315 320

His Glu Leu Thr Lys Lys Trp Leu Leu Glu Asp Phe Ser Cys Lys Ile
 325 330 335

Val Pro Tyr Ile Glu Val Ala Ser Leu Gly Val Thr Thr Phe Thr Leu
 340 345 350

Cys Ala Leu Cys Ile Asp Arg Phe Arg Ala Ala Thr Asn Val Gln Met
 355 360 365

Tyr Tyr Glu Met Ile Glu Asn Cys Ser Ser Thr Thr Ala Lys Leu Ala
 370 375 380

Val Ile Trp Val Gly Ala Leu Leu Leu Ala Leu Pro Glu Val Val Leu
 385 390 395 400

Arg Gln Leu Ser Lys Glu Asp Leu Gly Phe Ser Gly Arg Ala Pro Ala
 405 410 415

Glu Arg Cys Ile Ile Lys Ile Ser Pro Asp Leu Pro Asp Thr Ile Tyr
 420 425 430

Val Leu Ala Leu Thr Tyr Asp Ser Ala Arg Leu Trp Trp Tyr Phe Gly
 435 440 445

Cys Tyr Phe Cys Leu Pro Thr Leu Phe Thr Ile Thr Cys Ser Leu Val
 450 455 460

Thr Ala Arg Lys Ile Arg Lys Ala Glu Lys Ala Cys Thr Arg Gly Asn
 465 470 475 480

Lys Arg Gln Ile Gln Leu Glu Ser Gln Met Asn Cys Thr Val Val Ala
 485 490 495

Leu Thr Ile Leu Tyr Gly Phe Cys Ile Ile Pro Glu Asn Ile Cys Asn
 500 505 510

Ile Val Thr Ala Tyr Met Ala Thr Gly Val Ser Gln Gln Thr Met Asp
 515 520 525

Leu Leu Asn Ile Ile Ser Gln Phe Leu Leu Phe Phe Lys Ser Tyr Val
 530 535 540

Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe
 545 550 555 560

Met Glu Cys Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser

565

570

575

Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu
 580 585 590

Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser
 595 600 605

Val Gly Thr His Cys
 610

<210> 78
 <211> 1086
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 78
 atgtcccctg aatgcgcgcg ggcagcgggc gacgcgccct tgcgcagcct ggagcaagcc 60
 aaccgcaccc gctttccctt cttctccgac gtcaagggcg accaccggct ggtgctggcc 120
 gcggtggaga caaccgtgct ggtgctcatc tttgcagtgt cgctgctggg caacgtgtgc 180
 gccctggtgc tgggtggcgcg ccgacgacgc cgcggcgcca ctgcctgcct ggtactcaac 240
 ctcttctgcg cggacctgct cttcatcagc gctatccctc tgggtgctggc cgtgcgctgg 300
 actgaggcct ggctgctggg ccccgttgcc tgccacctgc tcttctacgt gatgaccctg 360
 agcggcagcg tcaccatcct cacgctggcc gcggtcagcc tggagcgcat ggtgtgcatc 420
 gtgcacctgc agcgcggcgt gcggggtcct gggcggcggg cgcgggcagt gctgctggcg 480
 ctcatctggg gctattcggc ggtcgccgct ctgcctctct gcgtcttctt tcgagtcgtc 540
 ccgcaacggc tccccggcgc cgaccaggaa atttcgattt gcacactgat ttggcccacc 600
 attcctggag agatctcgtg ggatgtctct tttgttactt tgaacttctt ggtgccagga 660
 ctggtcattg tgatcagtta ctccaaaatt ttacagatca caaaggcatc aaggaagagg 720
 ctcacggtaa gcctggccta ctcgagagc caccagatcc gcgtgtccca gcaggacttc 780
 cggctcttcc gcaccctctt cctcctcatg gtctccttct tcatcatgtg gagccccatc 840
 ttcataacca tctcctcat cctgatccag aacttcaagc aagacctggt catctggccg 900
 tccctcttct tctgggtggg ggccttcaca tttgctaatt cagccctaaa ccccatcctc 960
 tacaacatga cactgtgcag gaatgagtgg aagaaaattt tttgctgctt ctggttccca 1020

gaaaagggag ccattttaac agacacatct gtcaaaagaa atgacttgtc gattatttct 1080

ggctaa 1086

<210> 79
 <211> 361
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 79

Met Ser Pro Glu Cys Ala Arg Ala Ala Gly Asp Ala Pro Leu Arg Ser
 1 5 10 15

Leu Glu Gln Ala Asn Arg Thr Arg Phe Pro Phe Phe Ser Asp Val Lys
 20 25 30

Gly Asp His Arg Leu Val Leu Ala Ala Val Glu Thr Thr Val Leu Val
 35 40 45

Leu Ile Phe Ala Val Ser Leu Leu Gly Asn Val Cys Ala Leu Val Leu
 50 55 60

Val Ala Arg Arg Arg Arg Arg Gly Ala Thr Ala Cys Leu Val Leu Asn
 65 70 75 80

Leu Phe Cys Ala Asp Leu Leu Phe Ile Ser Ala Ile Pro Leu Val Leu
 85 90 95

Ala Val Arg Trp Thr Glu Ala Trp Leu Leu Gly Pro Val Ala Cys His
 100 105 110

Leu Leu Phe Tyr Val Met Thr Leu Ser Gly Ser Val Thr Ile Leu Thr
 115 120 125

Leu Ala Ala Val Ser Leu Glu Arg Met Val Cys Ile Val His Leu Gln
 130 135 140

Arg Gly Val Arg Gly Pro Gly Arg Arg Ala Arg Ala Val Leu Leu Ala
 145 150 155 160

Leu Ile Trp Gly Tyr Ser Ala Val Ala Ala Leu Pro Leu Cys Val Phe

165 170 175
 Phe Arg Val Val Pro Gln Arg Leu Pro Gly Ala Asp Gln Glu Ile Ser
 180 185 190
 Ile Cys Thr Leu Ile Trp Pro Thr Ile Pro Gly Glu Ile Ser Trp Asp
 195 200 205
 Val Ser Phe Val Thr Leu Asn Phe Leu Val Pro Gly Leu Val Ile Val
 210 215 220
 Ile Ser Tyr Ser Lys Ile Leu Gln Ile Thr Lys Ala Ser Arg Lys Arg
 225 230 235 240
 Leu Thr Val Ser Leu Ala Tyr Ser Glu Ser His Gln Ile Arg Val Ser
 245 250 255
 Gln Gln Asp Phe Arg Leu Phe Arg Thr Leu Phe Leu Leu Met Val Ser
 260 265 270
 Phe Phe Ile Met Trp Ser Pro Ile Phe Ile Thr Ile Leu Leu Ile Leu
 275 280 285
 Ile Gln Asn Phe Lys Gln Asp Leu Val Ile Trp Pro Ser Leu Phe Phe
 290 295 300
 Trp Val Val Ala Phe Thr Phe Ala Asn Ser Ala Leu Asn Pro Ile Leu
 305 310 315 320
 Tyr Asn Met Thr Leu Cys Arg Asn Glu Trp Lys Lys Ile Phe Cys Cys
 325 330 335
 Phe Trp Phe Pro Glu Lys Gly Ala Ile Leu Thr Asp Thr Ser Val Lys
 340 345 350
 Arg Asn Asp Leu Ser Ile Ile Ser Gly
 355 360

<210> 80
 <211> 1086
 <212> DNA
 <213> Unknown

<220>

<223> Novel Sequence

<400> 80

```

atgtcccctg aatgcgcgcg ggcagcgggc gacgcgccct tgcgcagcct ggagcaagcc      60
aaccgcaccc gctttccctt cttctccgac gtcaagggcg accaccggct ggtgctggcc      120
gcggtggaga caaccgtgct ggtgctcatc tttgcagtgt cgctgctggg caacgtgtgc      180
gccttggtgc tgggtggcgcg ccgacgacgc cgcggcgcga ctgcctgcct ggtactcaac      240
ctcttctgcy cggacctgct cttcatcagc gctatccctc tgggtgctggc cgtgcgctgg      300
actgaggcct ggctgctggg ccccggtgcc tgccacctgc tcttctacgt gatgaccctg      360
agcggcagcg tcaccatcct cacgctggcc gcggtcagcc tgaatcgcat ggtgtgcatc      420
gtgcacctgc agcgcggcgt gcggggctct gggcggcgcg cgcgggcagt gctgctggcg      480
ctcatctggg gctattcggc ggtcgccgct ctgcctctct gcgtcttctt tcgagtcgtc      540
ccgcaacggc tccccggcgc cgaccaggaa atttcgattt gcacactgat ttggcccacc      600
attcctggag agatctcgtg ggatgtctct tttgttactt tgaacttctt ggtgccagga      660
ctgggtcattg tgatcagtta ctccaaaatt ttacagatca caaaggcatc aaggaagagg      720
ctcacggtaa gcctggccta ctcggagagc caccagatcc gcgtgtccca gcaggacttc      780
cggtctttcc gcacctctt cctcctcatg gtctccttct tcatcatgtg gagccccatc      840
atcatcacca tcctcctcat cctgatccag aacttcaagc aagacctggt catctggccg      900
tccctcttct tctgggtggg ggccttcaca tttgctaatt cagccctaaa ccccatcctc      960
tacaacatga cactgtgcag gaatgagtgg aagaaaattt tttgctgctt ctggttccca     1020
gaaaagggag ccattttaac agacacatct gtcaaaagaa atgacttgct gattatttct     1080
ggctaa                                           1086

```

<210> 81

<211> 361

<212> PRT

<213> Unknown

<220>

<223> Novel Sequence

<400> 81

```

Met Ser Pro Glu Cys Ala Arg Ala Ala Gly Asp Ala Pro Leu Arg Ser
1           5           10           15

```

```

Leu Glu Gln Ala Asn Arg Thr Arg Phe Pro Phe Phe Ser Asp Val Lys

```

	20		25		30	
Gly Asp His Arg Leu Val Leu Ala Ala Val Glu Thr Thr Val Leu Val	35		40		45	
Leu Ile Phe Ala Val Ser Leu Leu Gly Asn Val Cys Ala Leu Val Leu	50		55		60	
Val Ala Arg Arg Arg Arg Arg Gly Ala Thr Ala Cys Leu Val Leu Asn	65		70		75	80
Leu Phe Cys Ala Asp Leu Leu Phe Ile Ser Ala Ile Pro Leu Val Leu		85		90		95
Ala Val Arg Trp Thr Glu Ala Trp Leu Leu Gly Pro Val Ala Cys His		100		105		110
Leu Leu Phe Tyr Val Met Thr Leu Ser Gly Ser Val Thr Ile Leu Thr		115		120		125
Leu Ala Ala Val Ser Leu Asn Arg Met Val Cys Ile Val His Leu Gln		130		135		140
Arg Gly Val Arg Gly Pro Gly Arg Arg Ala Arg Ala Val Leu Leu Ala	145		150		155	160
Leu Ile Trp Gly Tyr Ser Ala Val Ala Ala Leu Pro Leu Cys Val Phe		165		170		175
Phe Arg Val Val Pro Gln Arg Leu Pro Gly Ala Asp Gln Glu Ile Ser		180		185		190
Ile Cys Thr Leu Ile Trp Pro Thr Ile Pro Gly Glu Ile Ser Trp Asp		195		200		205
Val Ser Phe Val Thr Leu Asn Phe Leu Val Pro Gly Leu Val Ile Val		210		215		220
Ile Ser Tyr Ser Lys Ile Leu Gln Ile Thr Lys Ala Ser Arg Lys Arg	225		230		235	240
Leu Thr Val Ser Leu Ala Tyr Ser Glu Ser His Gln Ile Arg Val Ser		245		250		255

Gln Gln Asp Phe Arg Leu Phe Arg Thr Leu Phe Leu Leu Met Val Ser
 260 265 270

Phe Phe Ile Met Trp Ser Pro Ile Ile Ile Thr Ile Leu Leu Ile Leu
 275 280 285

Ile Gln Asn Phe Lys Gln Asp Leu Val Ile Trp Pro Ser Leu Phe Phe
 290 295 300

Trp Val Val Ala Phe Thr Phe Ala Asn Ser Ala Leu Asn Pro Ile Leu
 305 310 315 320

Tyr Asn Met Thr Leu Cys Arg Asn Glu Trp Lys Lys Ile Phe Cys Cys
 325 330 335

Phe Trp Phe Pro Glu Lys Gly Ala Ile Leu Thr Asp Thr Ser Val Lys
 340 345 350

Arg Asn Asp Leu Ser Ile Ile Ser Gly
 355 360

<210> 82
 <211> 1212
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 82
 atggcttgca atggcagtgc ggccaggggg cactttgacc ctgaggactt gaacctgact 60
 gacgaggcac tgagactcaa gtacctgggg ccccagcaga cagagctgtt catgcccata 120
 tgtgccacat acctgctgat ctctgtggtg ggcgctgtgg gcaatgggct gacctgtctg 180
 gtcatacctgc gccacaaggc catgcgcacg cctaccaact actacctctt cagcctggcc 240
 gtgtcgggacc tgctggtgct gctggtgggc ctgcccctgg agctctatga gatgtggcac 300
 aactaccctt tcctgctggg cggtggtggc tgctatttcc gcacgctact gtttgagatg 360
 gtctgcctgg cctcagtgtc caacgtcact gccctgagcg tggaacgcta tgtggccgtg 420
 gtgcacccac tccaggccag gtccatggtg acgcggggcc atgtgcgccg agtgcttggg 480
 gccgtctggg gtcttgccat gctctgtctc ctgcccaaca ccagcctgca cggcatccgg 540

```

cagctgcacg tgcctgccg gggcccagtg ccagactcag ctgtttgcat gctgggccgc      600
ccacggggccc tctacaacat ggtagtgacg accaccgcgc tgctcttctt ctgcctgccc      660
atggccatca tgagcgtgct ctacctgctc attgggctgc gactgcggcg ggagaggctg      720
ctgctcatgc aggaggccaa gggcaggggc tctgcagcag ccaggccag atacacctgc      780
aggctccagc agcacgatcg gggccggaga caagtgaaaa agatgctgtt tgtcctggtc      840
gtggtgtttg gcatctgctg ggccccgttc cagcccgacc gcgtcatgtg gagcgtcgtg      900
tcacagtga cagatggcct gcacctggcc ttccagcacg tgcacgtcat ctccggcatc      960
ttcttctacc tgggctcggc ggccaacccc gtgctctata gcctcatgtc cagccgcttc     1020
cgagagacct tccaggaggc cctgtgcctc ggggcctgct gccatcgct cagaccccg      1080
cacagctccc acagcctcag caggatgacc acaggcagca ccctgtgtga tgtgggctcc     1140
ctgggcagct ggggtccaccc cctggctggg aacgatggcc cagaggcgca gcaagagacc     1200
gatccatcct ga                                                                1212

```

<210> 83
 <211> 403
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 83

```

Met Ala Cys Asn Gly Ser Ala Ala Arg Gly His Phe Asp Pro Glu Asp
1           5           10           15

```

```

Leu Asn Leu Thr Asp Glu Ala Leu Arg Leu Lys Tyr Leu Gly Pro Gln
          20           25           30

```

```

Gln Thr Glu Leu Phe Met Pro Ile Cys Ala Thr Tyr Leu Leu Ile Phe
          35           40           45

```

```

Val Val Gly Ala Val Gly Asn Gly Leu Thr Cys Leu Val Ile Leu Arg
          50           55           60

```

```

His Lys Ala Met Arg Thr Pro Thr Asn Tyr Tyr Leu Phe Ser Leu Ala
          65           70           75           80

```

```

Val Ser Asp Leu Leu Val Leu Leu Val Gly Leu Pro Leu Glu Leu Tyr
          85           90           95

```

Glu Met Trp His Asn Tyr Pro Phe Leu Leu Gly Val Gly Gly Cys Tyr
 100 105 110

Phe Arg Thr Leu Leu Phe Glu Met Val Cys Leu Ala Ser Val Leu Asn
 115 120 125

Val Thr Ala Leu Ser Val Glu Arg Tyr Val Ala Val Val His Pro Leu
 130 135 140

Gln Ala Arg Ser Met Val Thr Arg Ala His Val Arg Arg Val Leu Gly
 145 150 155 160

Ala Val Trp Gly Leu Ala Met Leu Cys Ser Leu Pro Asn Thr Ser Leu
 165 170 175

His Gly Ile Arg Gln Leu His Val Pro Cys Arg Gly Pro Val Pro Asp
 180 185 190

Ser Ala Val Cys Met Leu Val Arg Pro Arg Ala Leu Tyr Asn Met Val
 195 200 205

Val Gln Thr Thr Ala Leu Leu Phe Phe Cys Leu Pro Met Ala Ile Met
 210 215 220

Ser Val Leu Tyr Leu Leu Ile Gly Leu Arg Leu Arg Arg Glu Arg Leu
 225 230 235 240

Leu Leu Met Gln Glu Ala Lys Gly Arg Gly Ser Ala Ala Ala Arg Ser
 245 250 255

Arg Tyr Thr Cys Arg Leu Gln Gln His Asp Arg Gly Arg Arg Gln Val
 260 265 270

Lys Lys Met Leu Phe Val Leu Val Val Val Phe Gly Ile Cys Trp Ala
 275 280 285

Pro Phe His Ala Asp Arg Val Met Trp Ser Val Val Ser Gln Trp Thr
 290 295 300

Asp Gly Leu His Leu Ala Phe Gln His Val His Val Ile Ser Gly Ile
 305 310 315 320

Asp Pro Ser

<400>	84
atgaatggca cctacaacac ctgtggctcc agcgacctca cctggccccc agcgatcaag	60
ctgggcttct acgcctaactt gggcgctctg ctggtgctag gcctgctgct caacagcctg	120
gcgctctggg tgttctgctg ccgcatgcag cagtggacgg agaccgcgat ctacatgacc	180
aacctggcgg tggccgacct ctgcctgctg tgcaaccttg ccttcgtgct gaactccctg	240
cgagacacct cagacacgcc gctgtgccag ctctcccagg gcattctacct gaccaacagg	300
tacatgagca tcagcctggg cacggccaac gccgtggacc gctatgtggc cgtagcggcac	360
ccgtgcgtg cccgcgggct gcggtcccc aggcaggctg cggccgtgtg cgcggtcctc	420
tgggtgctgg tcatcggctc cctggtggct cgctggctcc tggggattca ggagggcggc	480
ttctgcttca ggagcaccog gcacaatttc aaactccatgc ggttcccgt gctgggattc	540
tacctgcccc tggccgtggg ggtcttctgc tcctgaagg ttgtgactgc cctggcccag	600
agggcaccaca ccgacgtggg gcaggcagag gccaccocga aggctaaacg catggtctgg	660
qccaacctcc tqgtgttcgt ggtctgcttc ctgcccctgc acgtggggct gacagtgcgc	720

ctcgcagtggt gctggaacgc ctgtgccctc ctggagacga tccgtcgcgc cctgtacata 780
 accagcaagc tctcagatgc caactgctgc ctggacgcc tctgctacta ctacatggcc 840
 aaggagttcc aggaggcgtc tgcaactggcc gtggctcccc gtgctaaggc ccacaaaagc 900
 caggactctc tgtgcgtgac cctcgcctaa 930

<210> 85
 <211> 309
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 85

Met Asn Gly Thr Tyr Asn Thr Cys Gly Ser Ser Asp Leu Thr Trp Pro
 1 5 10 15

Pro Ala Ile Lys Leu Gly Phe Tyr Ala Tyr Leu Gly Val Leu Leu Val
 20 25 30

Leu Gly Leu Leu Leu Asn Ser Leu Ala Leu Trp Val Phe Cys Cys Arg
 35 40 45

Met Gln Gln Trp Thr Glu Thr Arg Ile Tyr Met Thr Asn Leu Ala Val
 50 55 60

Ala Asp Leu Cys Leu Leu Cys Thr Leu Pro Phe Val Leu His Ser Leu
 65 70 75 80

Arg Asp Thr Ser Asp Thr Pro Leu Cys Gln Leu Ser Gln Gly Ile Tyr
 85 90 95

Leu Thr Asn Arg Tyr Met Ser Ile Ser Leu Val Thr Ala Ile Ala Val
 100 105 110

Asp Arg Tyr Val Ala Val Arg His Pro Leu Arg Ala Arg Gly Leu Arg
 115 120 125

Ser Pro Arg Gln Ala Ala Ala Val Cys Ala Val Leu Trp Val Leu Val
 130 135 140

Ile Gly Ser Leu Val Ala Arg Trp Leu Leu Gly Ile Gln Glu Gly Gly

145 150 155 160
 Phe Cys Phe Arg Ser Thr Arg His Asn Phe Asn Ser Met Arg Phe Pro
 165 170 175
 Leu Leu Gly Phe Tyr Leu Pro Leu Ala Val Val Val Phe Cys Ser Leu
 180 185 190
 Lys Val Val Thr Ala Leu Ala Gln Arg Pro Pro Thr Asp Val Gly Gln
 195 200 205
 Ala Glu Ala Thr Arg Lys Ala Lys Arg Met Val Trp Ala Asn Leu Leu
 210 215 220
 Val Phe Val Val Cys Phe Leu Pro Leu His Val Gly Leu Thr Val Arg
 225 230 235 240
 Leu Ala Val Gly Trp Asn Ala Cys Ala Leu Leu Glu Thr Ile Arg Arg
 245 250 255
 Ala Leu Tyr Ile Thr Ser Lys Leu Ser Asp Ala Asn Cys Cys Leu Asp
 260 265 270
 Ala Ile Cys Tyr Tyr Tyr Met Ala Lys Glu Phe Gln Glu Ala Ser Ala
 275 280 285
 Leu Ala Val Ala Pro Arg Ala Lys Ala His Lys Ser Gln Asp Ser Leu
 290 295 300
 Cys Val Thr Leu Ala
 305

<210> 86
 <211> 1446
 <212> DNA
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 86
 atgcggtggc tgtggcccct ggctgtctct cttgctgtga ttttggctgt ggggctaagc 60
 aggggtctctg ggggtgcccc cctgcacctg ggcaggcaca gagccgagac ccaggagcag 120
 cagagccgat ccaagagggg caccgaggat gaggaggcca agggcgtgca gcagtatgtg 180

```

cctgaggagt gggcggagta cccccggccc attcaccctg ctggcctgca gccaaccaag 240
cccttggtgg ccaccagccc taaccccgac aaggatgggg gcacccaga cagtgggcag 300
gaactgaggg gcaatctgac aggggcacca gggcagaggc tacagatcca gaaccccctg 360
tatccggtga ccgagagctc ctacagtgcc tatgccatca tgcttctggc gctggtggtg 420
tttgcggtgg gcattgtggg caacctgtcg gtcattgtgca tcgtgtggca cagctactac 480
ctgaagagcg cctggaactc catccttgcc agcctggccc tctgggattt tctggtcctc 540
tttttctgcc tccctattgt catcttcaac gagatcacca agcagaggct actgggtgac 600
gtttcttgte gtgccgtgcc cttcatggag gtctcctctc tgggagtcac gactttcagc 660
ctctgtgccc tgggcattga ccgcttccac gtggccacca gcaccctgcc caaggtaggg 720
cccatcgagc ggtgccaatc catcctggcc aagttggctg tcatctgggt gggctccatg 780
acgttggtg tgctgtgagc cctgctgtgg cagctggcac aggagcctgc cccaccatg 840
ggcacccctg actcatgcat catgaaacct tcagccagcc tgcccagatc cctgtattca 900
ctggtgatga cctaccagaa cgcccgatg tgggtgtact ttggctgcta cttctgcctg 960
cccatcctct tcacagtcac ctgccagctg gtgacatggc ggggtgcgagg cctccaggg 1020
aggaagtcag agtgcagggc cagcaagcac gagcagtgtg agagccagct caagagcacc 1080
gtggtgggcc tgaccgtggt ctacgccttc tgcaccctcc cagagaacgt ctgcaacatc 1140
gtggtggcct acctctccac cgagctgacc cgccagaccc tggacctcct gggcctcatc 1200
aaccagttct ccaccttctt caagggcgcc atcacccag tgctgctcct ttgcatctgc 1260
aggccgctgg gccaggcctt cctggactgc tgctgctgct gctgctgtga ggagtgcggc 1320
ggggcttcgg aggcctctgc tgccaatggg tcggacaaca agctcaagac cgaggtgtcc 1380
tcttccatct acttccacaa gcccaggag tcacccccac tctgcccct gggcacacct 1440
tgctga 1446

```

<210> 87
<211> 481
<212> PRT
<213> Unknown

<220>
<223> Novel Sequence

<400> 87

Met Arg Trp Leu Trp Pro Leu Ala Val Ser Leu Ala Val Ile Leu Ala

1	5	10	15
Val Gly Leu Ser Arg Val Ser Gly Gly Ala Pro Leu His Leu Gly Arg	20	25	30
His Arg Ala Glu Thr Gln Glu Gln Gln Ser Arg Ser Lys Arg Gly Thr	35	40	45
Glu Asp Glu Glu Ala Lys Gly Val Gln Gln Tyr Val Pro Glu Glu Trp	50	55	60
Ala Glu Tyr Pro Arg Pro Ile His Pro Ala Gly Leu Gln Pro Thr Lys	65	70	75
Pro Leu Val Ala Thr Ser Pro Asn Pro Asp Lys Asp Gly Gly Thr Pro	85	90	95
Asp Ser Gly Gln Glu Leu Arg Gly Asn Leu Thr Gly Ala Pro Gly Gln	100	105	110
Arg Leu Gln Ile Gln Asn Pro Leu Tyr Pro Val Thr Glu Ser Ser Tyr	115	120	125
Ser Ala Tyr Ala Ile Met Leu Leu Ala Leu Val Val Phe Ala Val Gly	130	135	140
Ile Val Gly Asn Leu Ser Val Met Cys Ile Val Trp His Ser Tyr Tyr	145	150	155
Leu Lys Ser Ala Trp Asn Ser Ile Leu Ala Ser Leu Ala Leu Trp Asp	165	170	175
Phe Leu Val Leu Phe Phe Cys Leu Pro Ile Val Ile Phe Asn Glu Ile	180	185	190
Thr Lys Gln Arg Leu Leu Gly Asp Val Ser Cys Arg Ala Val Pro Phe	195	200	205
Met Glu Val Ser Ser Leu Gly Val Thr Thr Phe Ser Leu Cys Ala Leu	210	215	220
Gly Ile Asp Arg Phe His Val Ala Thr Ser Thr Leu Pro Lys Val Arg	225	230	235
			240

Pro Ile Glu Arg Cys Gln Ser Ile Leu Ala Lys Leu Ala Val Ile Trp
 245 250 255

Val Gly Ser Met Thr Leu Ala Val Pro Glu Leu Leu Leu Trp Gln Leu
 260 265 270

Ala Gln Glu Pro Ala Pro Thr Met Gly Thr Leu Asp Ser Cys Ile Met
 275 280 285

Lys Pro Ser Ala Ser Leu Pro Glu Ser Leu Tyr Ser Leu Val Met Thr
 290 295 300

Tyr Gln Asn Ala Arg Met Trp Trp Tyr Phe Gly Cys Tyr Phe Cys Leu
 305 310 315 320

Pro Ile Leu Phe Thr Val Thr Cys Gln Leu Val Thr Trp Arg Val Arg
 325 330 335

Gly Pro Pro Gly Arg Lys Ser Glu Cys Arg Ala Ser Lys His Glu Gln
 340 345 350

Cys Glu Ser Gln Leu Lys Ser Thr Val Val Gly Leu Thr Val Val Tyr
 355 360 365

Ala Phe Cys Thr Leu Pro Glu Asn Val Cys Asn Ile Val Val Ala Tyr
 370 375 380

Leu Ser Thr Glu Leu Thr Arg Gln Thr Leu Asp Leu Leu Gly Leu Ile
 385 390 395 400

Asn Gln Phe Ser Thr Phe Phe Lys Gly Ala Ile Thr Pro Val Leu Leu
 405 410 415

Leu Cys Ile Cys Arg Pro Leu Gly Gln Ala Phe Leu Asp Cys Cys Cys
 420 425 430

Cys Cys Cys Cys Glu Glu Cys Gly Gly Ala Ser Glu Ala Ser Ala Ala
 435 440 445

Asn Gly Ser Asp Asn Lys Leu Lys Thr Glu Val Ser Ser Ser Ile Tyr
 450 455 460

Phe His Lys Pro Arg Glu Ser Pro Pro Leu Leu Pro Leu Gly Thr Pro
465 470 475 480

Cys

<210> 88
<211> 6
<212> PRT
<213> Unknown

<220>
<223> Novel Sequence

<400> 88

Thr Leu Glu Ser Ile Met
1 5

<210> 89
<211> 5
<212> PRT
<213> Unknown

<220>
<223> Novel Sequence

<400> 89

Glu Tyr Asn Leu Val
1 5

<210> 90
<211> 5
<212> PRT
<213> Unknown

<220>
<223> Novel Sequence

<400> 90

Asp Cys Gly Leu Phe
1 5

<210> 91
<211> 34
<212> DNA
<213> Unknown

<220>

<223> Novel Sequence

<400> 91
gatcaagctt ccatggcgtg ctgcctgagc gagg 34

<210> 92
<211> 53
<212> DNA
<213> Unknown

<220>

<223> Novel Sequence

<400> 92
gatcggtacc ttagaacagg ccgcagtcct tcaggttcag ctgcaggatg gtg 53

<210> 93
<211> 5
<212> PRT
<213> Unknown

<220>

<223> Novel Sequence

<400> 93

Gln Tyr Glu Leu Leu
1 5

<210> 94
<211> 5
<212> PRT
<213> Unknown

<220>

<223> Novel Sequence

<400> 94

Asp Cys Gly Leu Phe
1 5

<210> 95
<211> 1185
<212> DNA
<213> Unknown

<220>

<223> Novel Sequence

<400> 95
atgggctgcc tcggcaacag taagaccgag gaccagcgca acgaggagaa ggcgcagcgc 60

gagggccaaca aaaagatcga gaagcagctg cagaaggaca agcaggtcta ccggggccacg 120
 caccgcctgc tgctgctggg tgctggagag tctggcaaaa gcaccattgt gaagcagatg 180
 aggatcctac atgttaatgg gtttaacgga gagggcgggc aagaggaccc gcaggctgca 240
 aggagcaaca gcgatggtga gaaggccacc aaagtgcagg acatcaaaaa caacctgaag 300
 gagggcattg aaaccattgt ggccgccatg agcaacctgg tgcccccctg ggagctggcc 360
 aacctgaga accagttcag agtggactac attctgagcg tgatgaacgt gccaaacttt 420
 gacttcccac ctgaattota tgagcatgcc aaggctctgt gggaggatga gggagtctgt 480
 gcctgctacg agcgctcaa cgagtaccag ctgatcgact gtgcccagta cttcctggac 540
 aagattgatg tgatcaagca ggccgactac gtgccaagt accaggacct gcttcgctgc 600
 cgcgtcctga cctctggaat ctttgagacc aagtccagg tggacaaagt caacttccac 660
 atgttcgatg tgggcggcca gcgcgatgaa cgccgcaagt ggatccagt cttcaatgat 720
 gtgactgcca tcatcttctg ggtggccagc agcagctaca acatgggtcat ccgggaggac 780
 aaccagacca accgtctgca ggaggctctg aacctcttca agagcatctg gaacaacaga 840
 tggctgcgta ccatctctgt gatcctcttc ctcaacaagc aagatctgct tgctgagaag 900
 gtcctcgctg ggaaatcgaa gattgaggac tactttccag agttcgctcg ctacaccact 960
 cctgaggatg cgactcccga gcccgagag gaccacgcg tgacccgggc caagtacttc 1020
 atccgggatg agtttctgag aatcagcact gctagtggag atggacgtca ctactgctac 1080
 cctcacttta cctgcgccgt ggacactgag aacatccgcc gtgtcttcaa cgactgccgt 1140
 gacatcatcc agcgcatgca tcttcgcgac tgcgggctgt tttaa 1185

<210> 96
 <211> 393
 <212> PRT
 <213> Unknown

<220>
 <223> Novel Sequence

<400> 96

Met Gly Cys Leu Gly Asn Ser Lys Thr Glu Asp Gln Arg Asn Glu Glu
 1 5 10 15

Lys Ala Gln Arg Glu Ala Asn Lys Lys Ile Glu Lys Gln Leu Gln Lys
 20 25 30

Asp Lys Gln Val Tyr Arg Ala Thr His Arg Leu Leu Leu Leu Gly Ala
 35 40 45

Gly Glu Ser Gly Lys Ser Thr Ile Val Lys Gln Met Arg Ile Leu His
 50 55 60

Val Asn Gly Phe Asn Gly Glu Gly Gly Glu Glu Asp Pro Gln Ala Ala
 65 70 75 80

Arg Ser Asn Ser Asp Gly Glu Lys Ala Thr Lys Val Gln Asp Ile Lys
 85 90 95

Asn Asn Leu Lys Glu Ala Ile Glu Thr Ile Val Ala Ala Ser Asn Leu
 100 105 110

Val Pro Pro Val Glu Leu Ala Asn Pro Glu Asn Gln Phe Arg Val Asp
 115 120 125

Tyr Ile Leu Ser Val Met Asn Val Pro Asn Phe Asp Phe Pro Pro Glu
 130 135 140

Phe Tyr Glu His Ala Lys Ala Leu Trp Glu Asp Glu Gly Val Arg Ala
 145 150 155 160

Cys Tyr Glu Arg Ser Asn Glu Tyr Gln Leu Ile Asp Cys Ala Gln Tyr
 165 170 175

Phe Leu Asp Lys Ile Asp Val Ile Lys Gln Ala Asp Tyr Val Pro Ser
 180 185 190

Asp Gln Asp Leu Leu Arg Cys Arg Val Leu Thr Ser Gly Ile Phe Glu
 195 200 205

Thr Lys Phe Gln Val Asp Lys Val Asn Phe His Met Phe Asp Val Gly
 210 215 220

Gly Gln Arg Asp Glu Arg Arg Lys Trp Ile Gln Cys Phe Asn Asp Val
 225 230 235 240

Thr Ala Ile Ile Phe Val Val Ala Ser Ser Ser Tyr Asn Met Val Ile
 245 250 255

Arg Glu Asp Asn Gln Thr Asn Arg Leu Gln Glu Ala Leu Asn Leu Phe

260 265 270
 Lys Ser Ile Trp Asn Asn Arg Trp Leu Arg Thr Ile Ser Val Ile Leu
 275 280 285
 Phe Leu Asn Lys Gln Asp Leu Leu Ala Glu Lys Val Leu Ala Gly Lys
 290 295 300
 Ser Lys Ile Glu Asp Tyr Phe Pro Glu Phe Ala Arg Tyr Thr Thr Pro
 305 310 315 320
 Glu Asp Ala Thr Pro Glu Pro Gly Glu Asp Pro Arg Val Thr Arg Ala
 325 330 335
 Lys Tyr Phe Ile Arg Asp Glu Phe Leu Arg Ile Ser Thr Ala Ser Gly
 340 345 350
 Asp Gly Arg His Tyr Cys Tyr Pro His Phe Thr Cys Ala Val Asp Thr
 355 360 365
 Glu Asn Ile Arg Arg Val Phe Asn Asp Cys Arg Asp Ile Ile Gln Arg
 370 375 380
 Met His Leu Arg Asp Cys Gly Leu Phe
 385 390

<210> 97
 <211> 1014
 <212> DNA
 <213> Homo sapiens

<400> 97
 atgaactcgt gggacgcggg cctggcgggg ctactggtgg gcacgatggg cgtctcgtg 60
 ctgtccaacg cgctggtgct gctctgcctg ctgcacagcg cggacatccg ccgccaggcg 120
 ccggcgctct tcaccctgaa cctcacgtgc gggaacctgc tgtgcaccgt ggtcaacatg 180
 ccgctcacgc tggccggcgt cgtggcgagc cggcagccgg cgggcgaccg cctgtgccgc 240
 ctggctgcct tcctcgacac cttcctggct gccaaactcca tgctcagcat ggccgcgctc 300
 agcatcgacc gctgggtggc cgtggtcttc ccgctgagct accgggcca gatgccgcct 360
 ccgagatgcg cgctcatcct ggcttacacg tggctgcacg cgctcacctt ccagccgcgc 420
 gcgctcgccc tgtcctggct cggcttcac cagctgtacg cctcgtgcac gctgtgcagc 480

cggcggccgg acgagcgccg gcgcttcgcc gtattcactg gcgccttcca cgctctcagc 540
 ttctctgtct ccttcgtcgt gctctgtgc acgtacctca aggtgctcaa ggtggccgcg 600
 ttccattgca agcgcatcga cgtgatcacc atgcagacgc tcgtgctgct ggtggacctg 660
 caccacagtg tgcgggaacg ctgtctggag gagcagaagc ggaggcgaca gcgagccacc 720
 aagaagatca gcaccttcac agggaccttc cttgtgtgct tcgcgcccta tgtgatcacc 780
 aggctagtgg agctcttctc cacggtgccc atcggtctcc actggggggg gctgtccaag 840
 tgcttggcgt acagcaaggc cgcacccgac ccctttgtgt actccttact gcgacaccag 900
 taccgcaaaa gctgcaagga gattctgaac aggtctctgc acagacgctc catccactcc 960
 tctggcctca caggcgactc tcacagccag aacattctgc cgggtgtctga gtga 1014

<210> 98
 <211> 337
 <212> PRT
 <213> Homo sapiens

<400> 98

Met Asn Ser Trp Asp Ala Gly Leu Ala Gly Leu Leu Val Gly Thr Met
 1 5 10 15

Gly Val Ser Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Leu Leu His
 20 25 30

Ser Ala Asp Ile Arg Arg Gln Ala Pro Ala Leu Phe Thr Leu Asn Leu
 35 40 45

Thr Cys Gly Asn Leu Leu Cys Thr Val Val Asn Met Pro Leu Thr Leu
 50 55 60

Ala Gly Val Val Ala Gln Arg Gln Pro Ala Gly Asp Arg Leu Cys Arg
 65 70 75 80

Leu Ala Ala Phe Leu Asp Thr Phe Leu Ala Ala Asn Ser Met Leu Ser
 85 90 95

Met Ala Ala Leu Ser Ile Asp Arg Trp Val Ala Val Val Phe Pro Leu
 100 105 110

Ser Tyr Arg Ala Lys Met Pro Pro Pro Arg Cys Ala Leu Ile Leu Ala
 115 120 125

Tyr Thr Trp Leu His Ala Leu Thr Phe Pro Ala Ala Ala Leu Ala Leu
 130 135 140

Ser Trp Leu Gly Phe His Gln Leu Tyr Ala Ser Cys Thr Leu Cys Ser
 145 150 155 160

Arg Arg Pro Asp Glu Arg Leu Arg Phe Ala Val Phe Thr Gly Ala Phe
 165 170 175

His Ala Leu Ser Phe Leu Leu Ser Phe Val Val Leu Cys Cys Thr Tyr
 180 185 190

Leu Lys Val Leu Lys Val Ala Arg Phe His Cys Lys Arg Ile Asp Val
 195 200 205

Ile Thr Met Gln Thr Leu Val Leu Leu Val Asp Leu His Pro Ser Val
 210 215 220

Arg Glu Arg Cys Leu Glu Glu Gln Lys Arg Arg Arg Gln Arg Ala Thr
 225 230 235 240

Lys Lys Ile Ser Thr Phe Ile Gly Thr Phe Leu Val Cys Phe Ala Pro
 245 250 255

Tyr Val Ile Thr Arg Leu Val Glu Leu Phe Ser Thr Val Pro Ile Gly
 260 265 270

Ser His Trp Gly Val Leu Ser Lys Cys Leu Ala Tyr Ser Lys Ala Ala
 275 280 285

Ser Asp Pro Phe Val Tyr Ser Leu Leu Arg His Gln Tyr Arg Lys Ser
 290 295 300

Cys Lys Glu Ile Leu Asn Arg Leu Leu His Arg Arg Ser Ile His Ser
 305 310 315 320

Ser Gly Leu Thr Gly Asp Ser His Ser Gln Asn Ile Leu Pro Val Ser
 325 330 335

Glu

<210> 99
<211> 21
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 99
cgagaaggtg ctcaaggtgg c

21

<210> 100
<211> 30
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 100
gagaagagct ccactagcct ggtgatcaca

30

<210> 101
<211> 36
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 101
gaattcatga actcgtggga cgcgggcctg gcgggc

36

<210> 102
<211> 32
<212> DNA
<213> Unknown

<220>
<223> Novel Sequence

<400> 102
ctcgagtcac tcagacaccg gcagaatgtt ct

32

THIS PAGE BLANK (USPTO)